

***Kvalitetsvurdering på den  
utdanningspolitiske dagsorden  
– på vei mot en markedsstyrt  
utdanningsideologi?***

**Marthe Amundsen**

**Hovedoppgave  
Profesjonsstudiet i pedagogikk  
Undervisning og ledelse**

**Våren 2004**

**Universitetet i Oslo  
Det utdanningsvitenskapelige fakultet  
Pedagogisk forskningsinstitutt**

## SAMMENDRAG AV HOVEDOPPGAVEN I PEDAGOGIKK

**TITTEL: KVALITETSVURDERING PÅ DEN  
UTDANNINGSPOLITISKE DAGSORDEN – PÅ VEI  
MOT EN MARKEDSSTYRT  
UTDANNINGSIDEOLOGI?**

**AV: AMUNDSEN, Marthe**

**EKSAMEN: Profesjonsstudiet i pedagogikk,  
Undervisning og ledelse**

**SEMESTER: Våren 2004**

**STIKKORD:**

- Kvalitetsvurdering på den utdanningspolitiske dagsorden
- Evaluerings- og organisasjonsteori
- Diskursanalyse

Denne oppgaven er en teoretisk basert analyse av en problemstilling som er knyttet til undersøkelse av kvalitetsvurdering på den utdanningspolitiske dagsorden.

Problemstillingen er delt i to:

1. Hvorfor har kvalitetsvurdering utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden?
2. Hvilken argumentasjon legges til grunn *for* kvalitetsvurdering på den utdanningspolitiske dagsorden?

Jeg har valgt å belyse problemstillingen med teorier og perspektiver både fra evalueringsfeltet, fra organisasjonsteori og fra det politiske felt. Man kan tolke dette som at oppgaven befinner seg i et skjæringspunkt mellom en teoretisk, vitenskapelig arena og en samfunnsmessig, politisk arena.

Man kan hevde at det i denne oppgaven er en nær sammenheng mellom teoretisk og metodisk tilnærming. Oppgaven baserer seg på et syn på sannhet, viten og virkelighet som kan beskrives med *meningskonstruksjon*. I denne sammenheng fokuseres det på et humanistisk og hermeneutisk utgangspunkt, en sosialkonstruktivistisk tilnærming og diskursteori. I tillegg presenteres symbol- og verktøyperspektivene som sentrale for å belyse oppgavens problemstilling.

Oppgavens kildemateriale varierer fra fagbøker, teoretiske avhandlinger og artikler, avisartikler og politiske dokumenter. Kildene kan sies å være med på å skape forståelse for et felt som hele tiden beveger seg mellom teori, vitenskap, samfunn og politikk. Dermed kan man muligens hevde at disse kildene er med på utvide forståelsen av evalueringsfeltet som både et teoretisk felt og et politisk-administrativt felt. I tillegg kan man si at *samfunnet*, eller offentligheten, er med på å forme et fokus på evaluering og tolkningen av ulike evalueringer.

Man bør tolke denne oppgaven som et forsøk på å utvide forståelsen for fenomenet kvalitetsvurdering og dets inntog på den utdanningspolitiske dagsorden. Det er et mål i seg selv for oppgaven å skape en forståelse for kvalitetsvurdering som et kompleks fenomen som ikke uproblematisk løser skolepolitiske oppgaver og utfordringer. Dermed kan man oppfatte kvalitetsvurdering som et fenomen som det er viktig å drøfte for å analysere et eventuelt kvalitetsvurderingssystemets muligheter og begrensninger.

Oppgaven skisserer den ”blå bølge” som betyningsfull for fokuset på kvalitetsvurdering på den utdanningspolitiske dagsorden. Den ”blå bølge” kan beskrives med New Public Management, en filosofi som har preget den offentlige sektoren i den vestlige verden siden 1980-tallet. I denne oppgaven blir filosofien beskrevet og diskutert. I tillegg drøftes tilhørende bevegelser som man kan se på som betydningsfulle for kvalitetsvurderingsfokuset. Disse kan også oppfattes som grunnleggende for den argumentasjonen som legges til grunn *for* kvalitetsvurdering på den utdanningspolitiske dagsorden.

I forlengelsen av problemstillingen kan man ane et spørsmål om institusjonalisering. Det er i denne oppgaven problematisk å si at kvalitetsvurdering er institusjonalisert som system. Dette er det for tidlig å snakke om. Dermed er det også vanskelig å si noe om hvilke diskurser som har vunnet hegemoni på den utdanningspolitiske dagsorden.

Avslutningsvis i oppgaven fremsettes spørsmålet: Hvilke diskurser vinner hegemoni? I tillegg blir det stilt spørsmål om man kan snakke om at den ”blå bølge” vinner hegemoni på den utdanningspolitiske dagsorden. Spørsmålene blir drøftet i forhold til St.meld. nr. 30 (2003-2004) *Kultur for læring*.

Med henvisning til stortingsmeldingen kan det være problematisk å hevde at det er NPM-filosofien og den blå bølge som er bærer av kvalitetsvurderingsfokuset. Man bør heller se på meldingen som representant for en fortsettende diskursiv kamp omkring kvalitetsvurdering, der det ennå ikke er noe hegemoni å snakke om.

## Forord

Arbeidet med denne hovedoppgaven har vært spennende. Spesielt utfordrende har det vært å jobbe med et så aktuelt tema som kvalitetsvurdering. I løpet av arbeidsperioden kom jeg stadig over avisdebatter, tv-debatter og politiske dokumenter som var med på å aktualisere min oppgave. Dette gav meg ekstra motivasjon på veien. Jeg avslutter oppgavearbeidet kun få uker etter at stortingsmeldingen *Kultur for læring* ble offentliggjort. En vesentlig del av denne kan sies å være utviklingen av et nasjonalt kvalitetsvurderingssystem, og kan dermed være med på å aktualisere denne oppgaven.

Først og fremst vil jeg takke Sigmund Lieberg for veiledning og gode diskusjoner.

Skoleetaten/Utdanningsetaten i Oslo, Avdeling for pedagogisk utvikling og kvalitet, har hele tiden i prosessen vært til stor hjelp. En spesiell takk til Elin Reite og KVA-gruppen.

Monika Borgan, Cecilie Jensen og Stine W. Skjæret fortjener en medalje for gjennomlesing av oppgaven og inspirerende diskusjoner.

Til slutt vil jeg takke mine medstudenter, og ikke minst Kjellerutvalget, for uttallige diskusjoner, pauser og hyggestunder.

Oslo, 10. mai 2004

Marthe Amundsen

# INNHOLDFORTEGNELSE

<b>1. INNLEDNING .....</b>	<b>1</b>
1.1. AKTUALISERING OG MOTIV .....	1
1.2. OPPGAVENS PROBLEMSTILLING .....	4
1.2.1. Mål for oppgaven .....	4
<b>2. METODISKE BETRAKTNINGER.....</b>	<b>6</b>
2.1. EN TEORETISK ANALYSE .....	6
2.1.1. Valg av teori og begrepsapparat .....	7
2.1.2. Oppgavens begrensninger.....	8
2.2. TEKSTTOLKNING OG MENINGSKONSTRUKSJON .....	9
2.3. INNSAMLING AV KILDER .....	10
2.4. KUNNSKAPSDANNELSE? .....	11
2.5. DET DISKURSIVE FELT.....	13
2.6. ET HUMANISTISK OG HERMENEUTISK UTGANGSPUNKT .....	14
2.7. DISKURSANALYSE SOM TEORI OG METODE .....	15
2.7.1. Diskursteori .....	18
2.8. SYMBOL- OG VERKTØYPERSPEKTIVET .....	22
<b>3. BEGREPSDRØFTING .....</b>	<b>26</b>
3.1. HVA ER <i>KVALITET</i> ? .....	26
3.1.1. Kvalitetsbegrepet i bruk på den utdanningspolitiske dagsorden .....	29
3.2. EVALUERINGSBEGREPET.....	30
3.3. RASJONALITET, LEGITIMITET OG INSTITUSJONALISERING .....	34
<b>4. UTVIKLINGSLINJER.....</b>	<b>37</b>
4.1. TEORETISKE UTVIKLINGSLINJER .....	39
4.1.1. Et nasjonalt bilde .....	43
4.2. UTVIKLING PÅ DEN UTDANNINGSPOLITISKE DAGSORDEN .....	47
4.2.1. Fra skoledirektørembeter til utdanningskontorer til ... ..	48
4.3. EVALUERINGENS RETORIKK – SENTRALISERING ELLER DESENTRALISERING? .....	54
<b>5. ARGUMENTER FOR KVALITETSVURDERING.....</b>	<b>61</b>
5.1. NEW PUBLIC MANAGEMENT-FILOSOFIEN .....	63
5.2. EVALUERINGSBØLGGEN.....	66
5.2.1. NPM, evaluering og den responsive stat .....	68
5.3. KVALITETSBØLGGEN .....	71
5.3.1. (E)quality .....	72
5.3.2. NPM, evaluering, kvalitet og den refleksive modernitet.....	73
5.4. INTERNASJONALISERING OG TEKNOLOGI .....	75
5.4.1. Globaliseringens paradoks.....	76
5.5. NPM OG UTVIKLINGEN AV VERKTØY FOR EFFEKTIVITET OG KVALITET: INDIKATORTENKNING .....	77
5.5.1. Standardisering .....	83
5.5.2. Benchmarking som strategi .....	85
5.6. NPM OG MODERNISERING .....	86
5.6.1. En rasjonell og kontrollerbar verden? .....	88

5.7. BEHOV FOR EN NY IDENTITET?.....	89
5.7.1. Identitet, ritual og/eller symbolsk fasade?.....	90
<b>6. HVORFOR KVALITETSVURDERING?.....</b>	<b>93</b>
6.1. NPM – EN ”PROTECTED DISCOURSE”?.....	97
<b>7. INSTITUSJONALISERING .....</b>	<b>103</b>
7.1. HVILKE DISKURSER VINNER HEGEMONI? .....	104
7.2. DEN BLÅ BØLGE INN PÅ DEN UTDANNINGSPOLITISKE DAGSORDEN?.....	105
7.3. ET KOMPLEKS FENOMEN .....	109
<b>LITTERATUR.....</b>	<b>111</b>
<b>WEB-ADRESSER .....</b>	<b>119</b>

# 1. INNLEDNING

”Evalueringer beskriver ikke bare virkeligheden. De former den.” (Dahler-Larsen & Krogstrup 2001a:243).

## 1.1. Aktualisering og motiv

Peter Dahler-Larsen starter boken *Den Rituelle Refleksjon* med å konstantere at vi lever i evalueringens tidsalder hvor alt skal evalueres (Dahler-Larsen 2001a:7). Denne boken vekket min interesse fra første stund, og var med på å forme min oppfatning av og forståelse for evalueringsfeltet i mine studier. Spesielt interessant er Dahler-Larsens erindringer om at selv om evalueringsfeltet har ekspandert, sitter han igjen med en oppfatning av at evalueringer er ”fattige”. Han mener vi er i ferd med å overfokusere på evalueringsfenomenet og at bruken og forståelsen av evaluering er i ferd med å bli triviell. Det er dette han mener at fører til en ”fattig” tenkning omkring evalueringer. Han skriver: ”Det er tit enten tale om en naiv, rationalistisk styringstænkning eller et kynisk politisk syn, hvor evalueringer fremstår som en simpel kamp om interesser. Det er brug for lidt mere sofistikerede forståelser af fænomenet.” (ibid). Dette sitatet sier noe om mine interesser for evalueringsfeltet. Jeg har i lang tid vært nysgjerrig på hvilke tendenser som viser seg i samfunnet når det gjelder bruken av dette fenomenet. Har Dahler-Larsen rett i sine betraktninger omkring bruken av evalueringsfenomenet? Er det kun snakk om *en simpel kamp om interesser*?

Nysgjerrigheten og interessen ble ytterligere vekket da jeg etter hvert ble kjent med organisasjonsforskning og organisasjonsteori. Jeg ble betatt av hvordan teorier innen organisasjonsfeltet beskriver ideer og oppskrifter *på reise* i organisasjonssamfunnet og hvordan disse kan være med på å utforme en organisasjons hverdag. Dette er en tilnærming som setter søkelyset på hvordan *ideer på reise* blir til rasjonaliserte symboler da de implementeres i en organisasjon (Røvik 1998).



Da jeg var i praksis i Skoleetaten i Oslo<sup>1</sup>, høsten 2002, arbeidet jeg hovedsakelig med deres helhetlige kvalitetsvurderingssystem og fikk mulighet til å kople disse problemstillingene og teoriene til praksis. Denne koplingen utviklet seg etter hvert til å bli en kjerne i mitt videre arbeid; jeg ble stilt overfor problemstillinger omkring *forholdet mellom politikk og administrasjon, arbeidet med et kvalitetsvurderingssystem, konsekvensene av et kvalitetsvurderingssystem og begrepers mening på ulike nivåer i en politisk-administrativ organisasjon*. Jeg begynte å reflektere over om Dahler-Larsen hadde rett; at bruken og forståelsen for evalueringsfenomenet var i ferd med å bli triviell. Spørsmål omkring hvorfor kvalitetsvurdering var et så sentralt element på den utdanningspolitiske dagsorden ble sentrale: Hva inneholder et kvalitetsvurderingssystem? Hvor kommer ideen om et kvalitetsvurderingssystem fra? Hva er hensikten med et kvalitetsvurderingssystem? Hvilke samfunnsinteresser skjuler seg bak et kvalitetsvurderingssystem?

Fokuset på kvalitetsvurdering ble for meg et viktig element i mine videre refleksjoner da jeg begynte å jobbe med denne hovedoppgaven. Jeg stod overfor mange spørsmål som jeg ville jobbe videre med. Spesielt ville jeg se nærmere på spørsmålet om hvorfor kvalitetsvurdering var blitt et sentralt element på den utdanningspolitiske dagsorden og hvordan arbeidet med dette foregikk. Jeg ville legge til grunn en historisk og samfunnsmessig diskusjon fordi disse kunne være med på å gi en utvidet forståelse av mine spørsmål. Det så ut som om diskusjonen omkring kvalitetsvurdering på den utdanningspolitiske dagsorden ikke var opptatt av mer bakenforliggende og grunnleggende problemstillinger som for eksempel hvilke konsekvenser dette kan få i et lengre tidsperspektiv, hvilke erfaringer som er gjort på liknende områder før, hvilke tendenser som influerer og er pådrivere, og hvordan kvalitetsvurdering i seg selv kan endre et system.

Det *historiske* viser her først og fremst til evalueringshistorie, mens *samfunnsmessig* sier noe om hvordan endringer i samfunnet, og organisasjonssamfunnet, kan være med på å påvirke, konstruere og ”presse” frem et fokus på kvalitetsvurdering på den utdanningspolitiske dagsorden. Sentralt i denne diskusjonen er å fremme en forståelse

---

<sup>1</sup> Skoleetaten i Oslo har skiftet navn til Utdanningsetaten i Oslo (mars 2004).

for hvorfor og hvordan dette skjer. Diskusjonen plasserer seg i et skjæringspunkt mellom vitenskap/teori og samfunn/politikk.

Foss Hansen understreker at det er viktig med en historisk analyse. Hun mener at diskusjonen om hvilke forhold som influerer utviklingen av evaluering nødvendigvis må baseres på en historisk analyse (Foss Hansen 2001). Hun ser at de fleste bidragene til evalueringshistorien er beskrivende. Dermed kan man si at det er få bidrag som viser en teoretisk og samfunnsanalytisk fundert forklaring relatert til sprednings- og utviklingsprosess, og som kan belyse mer grunnleggende spørsmål som hvorfor akkurat denne evalueringen ble aktuell og hvordan utviklingsprosessen artet seg. Foss Hansen er opptatt av dette perspektivet, og mener at litteraturen det her er snakk om har fokusert primært på beskrivelser av "... hvem der i relation til evaluering har vært aktive hvornår og hvordan." (op.cit:109). Hun mener en analyse av hvilke forhold som påvirker evalueringspraksisen, vil være et bidrag til å utvide historien.

Foss Hansen setter her ord på de problemstillinger som etter hvert utviklet seg som sentrale for meg. Den analysen hun beskriver kan være til hjelp for å utvide selve forståelsen av fenomenet *kvalitetsvurdering*. Man kan hevde at de bakenforliggende interessene ofte blir borte når man fremstiller evalueringer i rapporter; i stedet for å diskutere interessene, diskuterer man bare metoden. Dette feltet, interessefeltet, er spennende å løfte frem som premiss for diskusjonen om kvalitetsvurderingens inntog på den utdanningspolitiske dagsorden. Med en slik tilnærming er det sentralt å spørre hvilke meninger som vinner frem i det diskursive felt. I den sammenheng kan man hevde at det grunnleggende er å fremheve forholdet mellom samfunn/politikk og vitenskap/teori, og forsøke å finne frem til hvilke diskurser som vinner hegemoni på den utdanningspolitiske dagsorden.

Med en slik forståelse følger det at både en historisk og samfunnsmessig analyse er viktig for å besvare min problemstilling.

## 1.2. Oppgavens problemstilling

Oppgavens problemstilling er delt i to spørsmål:

1. Hvorfor har kvalitetsvurdering utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden?
2. Hvilken argumentasjon legges til grunn *for* kvalitetsvurdering på den utdanningspolitiske dagsorden?

*Den utdanningspolitiske dagsorden* viser til en samfunnsmessig åpen dagsorden som bæres av politiske aktører. I et demokratisk perspektiv er denne tilgjengelig for alle, blant annet gjennom media og de premissene som settes for Storting og regjering. Det er den allmenne politiske debatten om utdanning det her blir fokusert på.

*Kvalitetsvurdering* er ikke et enkelt fenomen å gi et meningsinnhold eller definere. Man kan forstå fenomenet på ulike måter. I forbindelse med denne oppgaven blir kvalitetsvurdering forstått som et fenomen som etter hvert får flere og flere elementer knyttet til seg, og som kan utvikle seg til å bli et system med form og innhold. Dermed kan man hevde at det utvikler seg til å bli et *kvalitetsvurderingssystem* som viser til en bestemt mening eller forståelse av dette systemet. Fenomenet har altså blitt satt i et system hvor ulike momenter<sup>2</sup> er med på å gi én bestemt mening.

En nærmere drøfting av begrepene ”kvalitet” og ”evaluering” finnes i kapittel 3.

### 1.2.1. Mål for oppgaven

Hovedmålet med oppgaven er å forsøke å bidra til å utvide forståelsen for fenomenet *kvalitetsvurdering* og dets inntog på den utdanningspolitiske dagsorden. Jeg vil prøve å vise at kvalitetsvurdering ikke er et enkelt fenomen som uproblematisk løser skolepolitiske oppgaver og utfordringer. Kvalitetsvurdering sees her som et kompleks fenomen som det er viktig å drøfte for å analysere et eventuelt kvalitetsvurderingssystems muligheter og begrensninger.

---

<sup>2</sup> Se kapittel 2.7.1 om momenter og elementer.

Jeg vil bruke en diskursteoretisk tilnærming og undersøke hvilke mønstre som finnes i aktuelle utsagn som knytter seg til kvalitetsvurdering, og hvilke sosiale konsekvenser ulike diskursive fremstillinger av virkelighet får.

## 2. METODISKE BETRAKTNINGER

I dette kapittelet vil jeg reflektere over min metode, det vil si beskrive min fremgangsmåte og oppgavetype. Kapittelet er viet mye plass da det omhandler grunnleggende refleksjoner som har betydning for å forstå min vinkling videre i oppgaven. I teksten vil det bli redegjort for hvilket syn på sannhet, viten og virkelighet som ligger til grunn for min egen tekstproduksjon og tolkning av andres tekstproduksjon. Mine metodiske betraktninger favner både fremgangsmåte og vitenskapelig ståsted da disse er problematisk å skille.

### 2.1. En teoretisk analyse

”Før jeg ved, hvad jeg skal undersøge, kan jeg ikke vide, hvordan jeg skal gjøre det.” (Holme & Solvang 1998:72<sup>3</sup>).

Denne oppgaven er en teoretisk basert analyse. Jeg bygger mine argumenter på teoretisk innsamlet materiale som er utviklet av andre teoretikere. Oppgaven er også samfunnsorientert og kan sies å bevege seg i et skjæringspunkt mellom en vitenskapelig arena, som viser til det teoretiske felt, og en samfunnsmessig arena, som viser til organisasjonssamfunnet og det politiske felt.

Valget av teoretisk oppgave var ikke klart for meg da jeg startet arbeidet med hovedoppgaven. Jeg var lenge sikker på at dette skulle bli en empirisk oppgave med fokus på forholdet mellom politikk og administrasjon. Temaet kvalitetsvurdering har ligget i bunn hele veien, men oppgaven har altså endret seg i arbeidsprosessen. Valget falt på teoretisk analyse fordi jeg etter hvert ble klar over hvor mange uklare og ubesvarte problemstillinger som fantes i feltet. Det ble riktig for meg å jobbe med et bidrag som kunne være med på å utvide forståelsen av fenomenet kvalitetsvurdering og dets inntog på den utdanningspolitiske dagsorden.

---

<sup>3</sup> Holme og Solvang refererer her til Jette Fog (1979).

### 2.1.1. Valg av teori og begrepsapparat

Et teoretisk studium impliserer en del valg, både valg av teori og begrepsapparat. Disse valgene er det jeg, som forfatter av oppgaven, som gjør. I tillegg er det viktig å reflektere over hvor litteraturen jeg velger kommer fra, det vil si hvem det er som har skrevet tekstene og hvilke vitenskapsteoretiske tradisjoner disse kommer fra. Mitt utvalg representerer bestemte tolkninger av temaet; den informasjonen, de ideene, dataene og teoriene jeg tilegner meg stammer fra teoretikernes egne ståsteder, noe de igjen har valgt for å oppnå én mulig forståelse. Tekstene jeg bruker inneholder derfor også valg, både valg av begreper, data og teori. I lys av dette kan man hevde at når man skal forfatte en teoretisk oppgave, handler det om å utvikle en bred forståelse for temaet sitt. Man bør orientere seg bredt i feltet, for å kunne gjøre de ”rette” valgene. *Kjennskap* til forskjellige tilnærminger og *oversikt* over feltet kan derfor oppfattes som grunnleggende i et teoretisk studium. Å studere forskningsfeltets *historie* kan være en god måte å oppnå slik forståelse på. Gjennom historisk lesning kan man utvikle en forståelse av hvordan feltet har utviklet seg, kanskje i forskjellige retninger, og hvem som har vært nøkkelteoretikere opp gjennom historien.

Når man skal skrive en hovedoppgave, har man en gylden mulighet til å fordype seg i sitt eget interessefelt. På den måten kan man utdype den generelle kunnskapen man har fra tidligere studier, og grave seg dypere ned i materien til det aktuelle, valgte feltet. Etter min mening er dette fundamentalt i en teoretisk analyse; det handler om å få grundig kjennskap til den eksisterende litteraturen i det valgte forskningsfeltet. Først da, når man er tilstrekkelig oppdatert på feltet, kan man skape *mening* i sin egen tekst og vurdere dennes gyldighet som teorigenerering.

I mitt forskningsprosjekt analyserer jeg mulige årsaker til hvorfor fenomenet *kvalitetsvurdering* har utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden, og hvilken argumentasjon som legges til grunn *for* kvalitetsvurdering her. En slik analyse er ikke mulig hvis den ikke relateres til et kriteriesett, eller kriterier for analysen (hvilke redskaper jeg bruker for å svare på forskningsspørsmålet mitt), og en programteori, som er mitt teoretiske bakteppe (det jeg tolker ut ifra). Man kan si at til sammen beskriver kriteriene og teorien hvilken *metode* som er brukt for å komme frem til plausible svar på problemstillingen. Mine kriterier for tolkning stammer fra en

humanistisk og sosialkonstruktivistisk virkelighetsforståelse, som kan oppfattes som grunnlaget for mitt vitenskapsteoretiske ståsted. Meningskonstruksjon kan sies å være en fellesnevner for det analysearbeidet jeg utfører. Her står diskursteori sentralt, i tillegg til mine analytiske innfallsvinkler til diskusjonen; verktøy- og symbolperspektivet. Dette er perspektiver som stammer fra nyinstitusjonell teori. I tillegg er det viktig for oppgavens helhetsforståelse å synliggjøre linjene mellom tematikken (evaluerings- og organisasjonsfeltet og politikk) og vitenskapsteorien. Det er dermed ikke bare teorier om kvalitetsvurdering som er sentrale i oppgaven, men også *hva* som gjør at dette fenomenet havner på den utdanningspolitiske dagsorden.

Valg av metodisk innfallsvinkel impliserer alltid konsekvenser som bør reflekteres over. Enten metodevalget er av kvantitativ eller kvalitativ art, kan det sies å være like viktig å kartlegge hvilke begrensninger som ligger i metodevalget, som det er å se hvilke muligheter dette valget gir. Denne refleksjonen er med på å sikre vitenskapelig gyldige (valide) og pålitelige (reliable) resultater<sup>4</sup>. Metodiske refleksjoner kan oppfattes som en viktig del av forskningsarbeidet som kan være med på å vise forskerens seriøsitet. Refleksjoner over metode kan dermed fortelle en eventuell leser om at forskeren har vært bevisst sine valg.

### 2.1.2. Oppgavens begrensninger

Min metodiske og vitenskapsteoretiske innfallsvinkel, valg av begreper og et metaperspektiv er med på å avgrense feltet. Slike begrensninger er naturlige og kan sies å være umulig å komme bort fra i en oppgave. *Avgrensning* i en oppgave er med på å sette fokus på bestemte elementer innenfor et felt, noe som gjør oppgaven mulig å forfatte og gir diskusjonen mulighet til å få en helhetlig form. Dette kan sees på som avgjørende for at teksten skal bli forståelig for leseren. Det er rimelig å hevde at oppgavens *begrensninger* henger sammen med oppgavens avgrensning og videre at dette er en viktig refleksjon man gjør seg underveis når man tar de avgjørende valgene. Det å forstå hvilke begrensninger som ligger i de valgene man tar, kan sies å være med på å avgrense oppgaven og utvikle dens gyldighet og pålitelighet.

---

<sup>4</sup> Se kapittel 2.4.

## 2.2. Teksttolkning og meningskonstruksjon

”Når jeg bruker et ord, betyr det nøyaktig det jeg velger at det skal bety – verken mer eller mindre.” (Caroll, L. 2000)<sup>5</sup>.

Forståelsen jeg sitter igjen med etter å ha studert en tekst, er påvirket av flere faktorer, slik som min profesjonelle bakgrunn og kvalifikasjoner, min *hensikt* med å lese de utvalgte tekstene, hva jeg *tror* den riktige tolkningen av tekstene er og hva jeg *tror* forfatteren av teksten vil si. Det bør sees på som en utopi at man kan forstå teksten slik forfatteren forsto den når den ble skrevet. Dette kan sies å være et kjent problem innen historisk teksttolkning, der man har gått fra å mene at teksten kan forstås nesten bedre enn forfatteren gjorde selv (ved nøye analyse) til å innse at det er mange influerende faktorer som spiller inn i tolkningsprosessen. Disse påvirkningene bør oppfattes som avgjørende for hvordan leseren til slutt forstår *meningen* med en tekst, det vil si at de fører til at jeg som leser kan tillegge argumentene en annen mening enn det som originalt var tekstens mening i forfatterens hode. Betti poengterer dette ved å skille mellom begrepene *meaning* og *significance* (Phillips 2000:36): *Meaning* viser til forfatterens intensjoner, mens *significance* peker på fortolkningen som leseren foretar på bakgrunn av sitt eget ståsted i sin egen kontekst. I tilknytning til dette kan man hevde at det er viktig å være klar over at enhver fortolkning av en tekst er situert i en gitt kontekst og at fortolkningen ikke nødvendigvis gir den samme forståelsen som forfatteren i sin opprinnelse intenderte å kommunisere.

Videre setter dette søkelyset på utfordringer en selv er midt i når man skal skrive en teoretisk oppgave. Jeg som fortolker av andres tekster, og forfatter av egen tekst på grunnlag av denne tolkningen, bør hele tiden være oppmerksom på de ulike fortolkningsmessige ”feller” jeg kan gå i underveis. Da er det ikke bare min egen tolkning av andres tekster jeg bør reflektere over, men også andres fortolkning av teorier og tekster. Det kan være mulig at forfatteren av den teksten jeg leser har mistolket både teorier og argumenter i sin egen tekstproduksjon.

---

<sup>5</sup> Sitatet er hentet fra en ny utgave av Lewis Carolls *Through the Looking-glass*, kapittel 6, hvor Humpty Dumpty samtaler med Alice (Alice i Eventyrland). Min oversettelse.



De fortolkningsmessige utfordringene som jeg har reflektert over ovenfor, kan beskrives nærmere med et sitat fra Ecos *The Limits of Interpretation*:

”... any act of interpretation is a dialectic between openness and form, initiative on the part of the interpreter and contextual pressure. ... Medieval interpreters were wrong in taking the world as a univocal text; modern interpreters are wrong in taking every text as an unshaped world. Texts are the human way to reduce the world to a manageable format, open to an intersubjective interpretive discourse. Which means that, when symbols are inserted into a text, there is, perhaps, no way to decide which interpretation is the “good” one, but it is still possible to decide, on the basis of the context, which one is due, not to an effort of understanding “that” text, but rather to hallucinatory response on the part of the addressee.” (Eco 1990:21).

### 2.3. Innsamling av kilder

Man kan si at teoretikere innenfor evalueringsfeltet kommer fra ulike forskningsfelt, som pedagogikk, sosiologi, statsvitenskap, samfunnsøkonomi, bedriftsøkonomi, som man videre kan hevde at representerer ulike vitenskapsteoretiske tilnærminger. Innen de forskjellige feltene har det blitt utviklet multiple metoder for evaluering. Derfor er det rimelig å anta at en evaluator/evalueringsteoretiker må orientere seg i flere felt og få kjennskap til ulike tilnærminger for å i det hele tatt ha mulighet til å si noe vitenskapsmessig holdbart om evaluanden<sup>6</sup>. I arbeidet med å finne relevante kilder, noe som har pågått under hele oppgaveprosessen, har jeg søkt ulike steder: databasen BIBSYS, kommunale, nasjonale og internasjonale nettsteder om utdanning, internasjonale nettsteder om evaluering og kvalitetsvurderingssystemer, tidsskrifter, aviser, medstudenter og faglig ansatte ved PFI (UiO), diverse kurs ved PFI/UV (UiO) og Skoleetaten i Oslo. Disse søkene har gitt meg tilgang til aktuell litteratur innen studiefeltet. Jeg har samlet litteratur fra ulike vitenskapsteoretiske retninger, jeg har fått tilgang på et stort dokumentmateriale (rammer, vedtak og utredninger), jeg har studert historie og tidligere forskning, i tillegg til metaforskning, komparativ forskning og teori-/metodebyggende forskning. Av dette følger det at kildene jeg har studert og bruker i oppgaven, er av ulik art. I tillegg til fagbøker, er mange av mine kilder hentet fra avhandlinger, artikkelsamlinger og tidsskrifter. Jeg har også studert politiske dokumenter som er å finne både i trykt utgave og på offentlige nettsider. Debattsider i diverse aviser kan også sees på som en del av mitt kildemateriale. Til sammen har dette

---

<sup>6</sup> Evaluanden er det som evalueres, enten det er en gruppe personer, et tiltak, et undervisningsopplegg osv. Evaluanden bør dermed ikke tolkes kun som en gjenstand, begrepet viser til det som evalueres.

gitt meg en relativt bred forståelse for evalueringsfeltet, ikke bare som et teoretisk felt, men også som et politisk-administrativt felt. Det politiske feltet bør derfor oppfattes som en del av mine kilder.

## 2.4. Kunnskapsdannelse?

Mitt teorivalg og begrepsapparat i oppgaven er ikke tilfeldig. Jeg har brukt lang tid på å komme frem til hvilke vitenskapsteoretiske perspektiver og tilhørende begreper som er hensiktsmessig å bruke for å belyse oppgavens problemstilling. Jeg satt med en del antagelser og påstander før jeg begynte å jobbe aktivt med hovedoppgaven, og jeg mener at jeg, etter å ha tilegnet meg relativt mye kunnskap omkring oppgavens tema, har kommet frem til både begreper og teorier som er legitime å bruke for å belyse problemstillingen og forsvare dette som teorigenerering. Holme og Solvang hevder at teori og teoriutvikling handler om å abstrahere et konkret fenomen og gjennom dette bli bedre i stand til å forstå, forklare og forutsi fenomenet (Holme & Solvang 1998:50). Jeg støtter denne forståelsen av teoriutvikling og mener jeg selv, med denne oppgaven, er med på å utvide forståelsen av fenomenet kvalitetsvurdering og dermed drive teoriutvikling. I oppgaven vil begrepene rasjonalitet og legitimitet, i tillegg til institusjonalisering, være sentrale. Jeg har valgt å bruke disse begrepene for å belyse hvordan *kvalitetsvurdering* kan være med på å bestemme den utdanningspolitiske dagsorden, hvilket bilde den gir av seg selv og hvilken retning den beveger seg i. Begrepene kan betraktes som sentrale i dagens forståelse av politikk og organisasjon, og kan derfor sies å være med på å belyse hvilke tendenser som påvirker den politiske dagsorden.

Begrepene jeg har valgt henger sammen med mitt syn på sannhet, viten og virkelighet. Begrepene er hele tiden i endring og bør derfor analyseres og diskuteres med en samtidshistorisk tilnærming. Det er etter min mening mulig å argumentere for at teorivalget i denne oppgaven passer sammen med en slik forståelse.

I min refleksjon omkring teoriutvikling og verifisering, og om jeg kan stå inne for at min oppgave kan være med på å skape en utvidet forståelse av fenomenet

kvalitetsvurdering, har jeg fundert på om jeg kan bruke vitenskapsmetodiske begreper som reliabilitet, validitet og generaliserbarhet. Min motstand mot å bruke disse begrepene henger sammen med at dette er begreper som stammer fra en annen vitenskapsforståelse enn min egen, nemlig fra et positivistisk forskningsparadigme, hvis hovedmålsetning kan sies å være å finne den ene objektive sannheten. Dermed kan man hevde at begrepene, i sin opprinnelige betydning, er uegnet innenfor forskning som anerkjenner at det finnes flere likeverdige sannheter. Imidlertid har Kvale (1997) og Østerud (1998) jobbet med å rekonseptualisere eller rekonstruere begrepene slik at de kan brukes innenfor kvalitativ forskning. Kvale hevder blant annet at reliabilitet kan forstås som *pålitelighet*, og validitet som *gyldighet* (Kvale 1997:160,164).

Innenfor et positivistisk forskningsparadigme hevdes det at det er bare gjennom høy reliabilitet at det er mulig å styrke funnenes validitet (Østerud 1998:122). Man kan hevde at denne regelen ikke vil gjelde i et sosialkonstruktivistisk forskningsparadigme fordi man har forkastet korrespondanseteoriens tro på at det kan eksistere en gjengivelse av virkeligheten som korresponderer i et "en til en"-forhold. Østerud sier at ingen virkelighetstolkninger blir sannere selv om de samsvarer med andre virkelighetstolkninger (op.cit:122-123). Man avstår fra å spørre om graden av overensstemmelse mellom teksten og virkeligheten, og spør i stedet om den vinner legitimitet hos en relevant gruppe forskere. Et sentralt poeng i Østeruds fremstilling er hvordan kunnskapsprodusenter forholder seg til de regler for kunnskapsfremstilling som gjelder for genren. Han belyser dette i følgende sitat: "Høy validitet eller troverdighet kan sies å være den "maske" en tekst ikler seg når den overbeviser den kompetente leser om at forskeren har overholdt genrens lover." (op.cit:125).

I en positivistisk tilnærming til validitet kan man anta at det er vanlig å spørre om man måler det man tror man måler. I et slikt perspektiv kan man dermed hevde at kvalitativ forskning blir ugyldig så lenge den ikke resulterer i tall, nettopp fordi det her er snakk om "måling". I en videre tolkning av validitet kan man omformulere spørsmålet til å gjelde i hvilken grad en metode undersøker det den er ment å undersøke, eller som Pervin har formulert det: "... i hvilken grad våre observasjoner faktisk reflekterer de fenomenene eller variablene som vi ønsker å vite noe om." (Kvale 1997:165). Med en slik forståelse av validitet kan man argumentere for at forskerens håndverksmessige

dyktighet blir sentral, og derigjennom at den kvalitative forskningen i prinsippet kan gi valide funn.

Konsekvensen av en slik forståelse av validitet blir i følge Kvale at forskeren må utsette den vitenskapelige kunnskapen som blir produsert for kontrollering, utspørring og teoretisering (op.cit:168-170). Med kontrollering mener han at forskeren har et kritisk syn på sine egne tolkninger, og spiller ”djevelens advokat” overfor sine egne funn. Utspørringskriteriet dreier seg om hvilke spørsmål som stilles til teksten, mens teoretisering kan knyttes til at forskeren må ha en teoretisk oppfatning av det fenomenet som undersøkes.

I spørsmålet om generaliserbarhet kan det hevdes at forskningsambisjonen i et sosialkonstruktivistisk perspektiv ikke er å avdekke lovmessigheter som kan generaliseres til andre kontekster, men å bidra til innsikt og forståelse for hvordan man konstruerer sosiale virkeligheter ved hjelp av (ulike) forestillinger om verden. Forestillinger om verden kan sies å interagere med større sosiale praksiser i en sosial prosess. Østerud belyser dette spørsmålet om generaliserbarhet i følgende sitat:

”Fra tradisjonelt positivistisk hold blir det gjerne innvendt mot undersøkelser som stanser før de har foretatt en slik ytre validering, at det er vanskelig å vurdere deres verdi. Omvendt rettes det innvendinger fra kritisk og konstruktivistisk orienterte forskere mot denne form for generalisering for at den er overflatisk. Den tar ikke hensyn til det forhold at konteksten er av avgjørende betydning for hvert enkelt av de fenomenene vi utforsker. Disse forskerne nøyer seg ofte med det man kan kalle en teoretisk generalisering.” (Østerud 1998:128).

## 2.5. Det diskursive felt

Oppgaven beveger seg hele tiden i skjæringspunktet mellom en teoretisk, vitenskapelig arena og en samfunnsmessig, politisk arena. Det er spenningene som ligger i dette skjæringspunktet, eller *det diskursive felt*, jeg ønsker å belyse da jeg vil undersøke *hvorfor kvalitetsvurdering* har utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden, og *hvilken argumentasjon* som legges til grunn for kvalitetsvurdering på den utdanningspolitiske dagsorden.

I dette skjæringspunktet kan man snakke om at det finnes en meningskamp, eller en diskursiv kamp, som gjør seg gjeldende. Denne kampen kan oppfattes som en kamp om hvilke meninger, det vil si hvilken politisk agenda, som skal være gjeldende og som prøver å få gjennomslag både hos profesjonelle og i offentligheten. Man kan hevde at det er dette som er det grunnleggende i diskursteorien, nemlig at det er i diskursene mening konstrueres i den sosiale verden, og forskjellige diskurser kjemper hele tiden mot hverandre for å oppnå hegemoni<sup>7</sup>.

Videre vil jeg beskrive mitt grunnlag for å tolke dette skjæringspunktet. Beskrivelsen bør sees på som et uttrykk for både mitt vitenskapsteoretiske ståsted, mitt syn på sannhet, viten og virkelighet, og min metode for å analysere og belyse oppgavens problemstilling.

## 2.6. Et humanistisk og hermeneutisk utgangspunkt

Mitt syn på sannhet, viten og virkelig kan sies å ha et humanistisk og hermeneutisk utgangspunkt. Historisk sett kan man hevde at min metode for å analysere og belyse oppgavens problemstilling stammer fra en hermeneutisk tradisjon.

Man kan argumentere for at åndsvitenskap, heriblant samfunnsvitenskap, har et annet erkjennelsesteoretisk grunnlag enn naturvitenskap. Debatten mellom ånds- og naturvitenskap foregikk blant annet i Tyskland på 1800-tallet (Ryen 2002). Den åndsvitenskapelige fløyen mente at deres vitenskap var ideografisk, det vil si at den viser til noe individuelt eller enkeltstående, og dermed ikke det nomotetiske eller lovmessige (op.cit:37). I tillegg kan man si de vektla at åndsvitenskap bygger på hermeneutisk tolkning og innleving. Hermeneutikk kan betraktes som tolkningens kunst og vitenskap. Ryen mener at denne retningen står i opposisjon til den holdningen som argumenterer for at man uforutinntatt kan forholde seg objektivt til såkalte rene fakta. Hun sier i tilknytning til dette: "Enhver forståelse er *qua* situasjonsbetinget nødvendigvis preget av allerede ervervede eller tilvante oppfatninger, altså *for-dommer* i en ikke nedsettende betydning av ordet." (op.cit:37).

---

<sup>7</sup> Diskursteorien forklares mer grunnleggende i kapittel 2.7/2.7.1.

Hermeneutikken danner grunnlaget for den *tolkende sosiologi*, som vi blant annet kan finne igjen hos Habermas og Giddens. Her søker man tolkninger av samfunnsmedlemmenes virksomhet, ikke kausalforklaringer. Handling er flertydig og tillater derfor flere tolkninger (Ryen 2002).

Oppsummerende kan man henvise til Eco for å beskrive hovedtrekkene i den hermeneutiske tradisjonen. Han viser til følgende fem punkter for å beskrive hva tradisjonen i hovedtrekk står for (Eco 1990:19-20):

1. En avvisning av målbare systemer. Dette viser til troen på at ingenting er stabilt og at alle elementer i universet er avhengig av hverandre.
2. En avvisning av kausalitetsprinsippet. Elementenes gjensidige handlinger følger ikke en lineær sekvens, men heller en spiralliknende logikk av gjensidige deltakende elementer.
3. En avvisning av dualisme, det vil si at den hermeneutiske tradisjonen ikke skiller kategorisk mellom for eksempel ånd og materie.
4. En avvisning av agnostisisme, det vil si at hermeneutikere avviser den oppfatning om at det ikke er mulig å erkjenne det oversanselige, det som går over all erfaring.
5. Den hermeneutiske tradisjonen baseres på at alt er kontekstuell og må derfor knyttes til en sammenheng.

## 2.7. Diskursanalyse som teori og metode

”Virkeligheden er samfundsskabt, og videnssociologien må analysere de processer, hvorunder denne skabelse finder sted.” (Berger & Luckmann 1992:13).

Man kan si at diskursanalysen har et sosialkonstruktivistisk utgangspunkt i sitt syn på språk og individforståelse. Det sosialkonstruktivistiske utgangspunktet kan man blant annet se i: (1) en kritisk innstilling overfor selvfølgelig viten, (2) historiske og kulturelle betingelser, (3) sammenhengen mellom viten og sosiale prosesser, og (4)

sammenhengen mellom viten og sosial handling (Winther Jørgensen & Phillips 1999:13-14).

Innenfor det sosialkonstruktivistiske virkelighetsperspektivet kan man se en erkjennelse av at mennesker gis muligheten til å påvirke og endre sin sosiale virkelighet. Videre kan man da si at fremtiden fremstår som noe som blir skapt av mennesker og samfunnsutviklingen, og dermed som noe som kan endres. Mennesket betraktes som handlingsagent i sin egen virkelighet.

Denne virkelighetsanskuelsen kan man finne igjen i Berger og Luckmanns teori om hvordan mennesker konstruerer virkeligheter (Berger & Luckmann 1992<sup>8</sup>). Teorien bygger på ideen om at virkelighet er en kvalitet vi mennesker tilskriver fenomenene gjennom sosial samhandling (ibid). En konsekvens av dette kan da være at det i verden ikke finnes én virkelighet, men mange forskjellige. Berger og Luckmann mener at virkeligheten skapes gjennom sosial samhandling, og følgelig at mennesker ikke skaper virkelighet i et sosialt vakuum. Gjennom sosial samhandling skapes sosiale normer, lover, vaner og antagelser. Disse prosessene pågår kontinuerlig. Gjennom prosessene konstrueres og konstitueres samfunnet som objektiv virkelighet: det skjer en institusjonaliseringsprosess (ibid).

Videre hevder Berger og Luckmann at menneskelig samhandling for det meste foregår gjennom språk, og på den måten konstrueres og konstitueres virkeligheter gjennom språkbruk. I motsetning til dette kan man muligens finne et syn på virkeligheten som oppfatter den som en størrelse som eksisterer uavhengig av språkbruk og sosial samhandling. Virkeligheten blir i denne sammenheng noe som finnes ”der ute” som en gitt størrelse, og språket kan i dette perspektivet oppfattes som om det korresponderer i et ”en til en”-forhold med objekter og hendelser i ”virkeligheten”. Dette perspektivet opererer med en distinksjon mellom den objektive virkelighet og den subjektive, oppfattede verden. Språket betraktes som en pålitelig og presis forbindelseslinje mellom det objektive og det subjektive (ibid). Det er ut fra Berger og Luckmanns teori rimelig å hevde at det sosialkonstruktivistiske perspektivet forkaster disse antagelsene og snur

---

<sup>8</sup> 1992-utgaven som det her refereres til er en dansk oversettelse. Boken kom for første gang ut i 1966.

dem på hodet ved å rette fokus mot hvordan språket vi bruker konstituerer våre virkeligheter.

Disse premissene kan betraktes som et utgangspunkt for diskursanalytiske tilnærminger i forskningssammenheng. Winther Jørgensen og Phillips fremhever at man ikke kan bruke diskursanalyse som analysemetode løsrevet fra det teoretiske og metodiske grunnlaget i ens forskning, men at de diskursanalytiske tilnærmingene de presenterer utgjør et teoretisk og metodisk hele. Diskursanalyse og teori er kjedet sammen, og man må akseptere de ontologiske og epistemologiske premissene som ligger til grunn for språkets rolle i den sosiale konstruksjonen av verden (Winther Jørgensen & Phillips 1999).

Diskursanalysen kan sees på som en leverandør av redskaper til å undersøke hvordan mening skapes sosialt, og hvilke virkninger bestemte verdensbilder har for våre sosiale liv. Begrepet *diskurs* beskriver en bestemt måte å snakke om og forstå, verden, eller et utsnitt av verden, på (op.cit:9). Man kan finne flere retninger innenfor diskursanalysen, blant annet diskursteori, kritisk diskursanalyse og diskurspsykologien<sup>9</sup> (op.cit:9-10). Her vil jeg gå spesielt inn på en av retningene, diskursteori. Diskursteori kan oppfattes som den retningen innen diskursanalyse som har teoriutvikling som mål og som vektlegger den diskursive kamp. Først vil jeg redegjøre for de grunnleggende begrepene i diskursanalysen.

I følge Winther Jørgensen og Phillips skiller diskursanalysen seg grovt sett fra andre tekstanalytiske innfallsvinkler ved at den vektlegger fenomenene *funksjon*, *konstruksjon* av virkelighet og *variasjon* i de ulike tekstenes fremstilling av det samme fenomenet (op.cit:18). Ordet funksjon kan sies å ha to hovedbetydninger: (1) den virksomhet eller handling en "ting" utfører for å oppfylle sin hensikt i et større system (språkvitenskap), og (2) forholdet mellom to eller flere variabler (matematikk). Begrepet konstruksjon viser at et sosialkonstruktivistisk perspektiv ligger til grunn, og variasjon spør hvordan fremstillingen av et fenomen kan variere avhengig av avsenders og mottakers forhold til fenomenet.

---

<sup>9</sup> Diskursanalyse favner over et stort antall teorier og metoder. Diskursanalyse er derfor ikke én tilnærming, men innebefatter mange tverrfaglige og multidisiplinære tilnærminger som man kan anvende på mange ulike sosiale områder i mange typer undersøkelser (Winther Jørgensen & Phillips 1999:9).



Diskursanalytikerens mål er ikke å komme ”bakom” diskursen i sine analyser. Man skal verken finne ut hva folk *virkelig* mener når de sier det ene eller det andre, eller finne ut hvordan virkeligheten *egentlig* er bak diskursen. Utgangspunktet, som vist ovenfor, er at man aldri kan nå virkeligheten utenom diskursene, og at det derfor er diskursen i seg selv som utgjør gjenstanden for analysen (op.cit:31). I lys av dette kan man hevde at det i den diskursanalytiske undersøkelse ikke først og fremst gjelder å sortere hvilke utsagn om verden som er riktige eller gale, men å undersøke hvilke mønstre som finnes i utsagnene, og hvilke sosiale konsekvenser forskjellige diskursive fremstillinger av virkeligheten får. Et sentralt problem i denne undersøkelsen, er at det alltid er vanskelig å undersøke de diskursene man selv er tett innpå og kan være forutinntatt overfor.

Man kan ut fra dette si at diskursanalysens hovedformål er å kartlegge de prosesser hvor det utspilles en ”kamp” om hvordan tegnenes betydning skal fastsettes eller bestemmes, og hvor noen meningsfikseringer blir så konvensjonaliserende at vi oppfatter dem som naturlige.

### 2.7.1. Diskursteori

Laclau og Mouffes diskursteori har sitt utgangspunkt i poststrukturalismen der diskursene konstruerer mening i den sosiale verden (Winther Jørgensen & Phillips 1999). Teorien tar utgangspunkt i at mening aldri kan fastlåses på grunn av språkets grunnleggende ustabilitet (op.cit:15). Ingen diskurs er dermed en lukket enhet, diskursen omformes hele tiden siden den alltid er i kontakt med andre diskurser. Derfor er nøkkelordet i teorien *diskursiv kamp*. Forskjellige diskurser, som hver for seg representerer en bestemt måte å snakke på og forstå den sosiale verden på, kjemper hele tiden mot hverandre for å oppnå hegemoni, det vil si de prøver å fastlåse språkets betydninger på nettopp deres måte (ibid).

I den diskursteoretiske tilnærmingen sees alle sosiale praksiser som diskursive (op.cit:29). Her skilles det ikke mellom diskursive og ikke-diskursive sosiale praksiser. Dette medfører at det ikke bare er tekst og tale som finnes, men at selve diskursen er

materiell og at for eksempel infrastruktur, institusjoner og økonomi oppfattes som former for diskurs. Det finnes derfor ikke et dialektisk samspill mellom diskurs og ”noe annet”: Diskursen er fullt ut konstituerende for vår verden (ibid).

Laclau og Mouffes fremsetter fire begreper som svært viktige i diskursteorien (op.cit:36). Det første begrepet er *diskurs* og kan forstås som en fastsettelse av mening innenfor et bestemt domene. *Momenter* er det andre sentrale begrepet. Alle tegnene i en diskurs er momenter; de er knuter på et fiskenett, og deres betydning blir holdt fast ved at de er forskjellige fra hverandre på bestemte måter (differential positions). Samtidig er momentene avhengig av hverandre. Hver for seg er tegnenes betydning bestemt av deres relasjoner til hverandre. Blant momentene peker det seg ut privilegerte tegn som kalles nodalpunkter<sup>10</sup>. Dette er tegn som står i sentrum, det vil si de tegn som en diskurs er organisert rundt.

Diskursteoriens tredje begrep er *elementer*. For å forstå hva som menes med elementer, må det først skapes en forståelse for ”det diskursive felt”. En diskurs er en reduksjon av muligheter, det betyr at diskursen i seg selv er et forsøk på å stanse tegnenes bevegelse i forhold til hverandre, og dermed er et forsøk på å skape entydighet. Alle de mulighetene som diskursen dermed utelukker, kaller Laclau og Mouffes for ”det diskursive felt” (op.cit:38). Dette er diskursens ytre, og kan betraktes som det man ignorerer for å skape mening, men som kan fremkomme i andre liknende diskurser. Nettopp fordi en diskurs alltid konstitueres i forhold til et slikt ytre, står den alltid i fare for å bli undergravd av dette ytre, altså for at andre definisjoner av tegnene skal forstyrre dens entydighet. Det er her begrepet *element* gjør seg gjeldende. Elementer er de tegnene som ikke endelig har fått en bestemt mening, det vil si de tegn som ennå er flertydige. Diskursen forsøker hele tiden å gjøre elementene om til momenter ved å redusere deres flertydighet til entydighet. Til slutt skisseres begrepet *artikulasjon* som viktig innen diskursteorien. Dette begrepet sier noe om betydningen av kontekst Laclau og Mouffes mener at tegn alltid må stå i relasjon til andre tegn for å skape mening. Artikulasjon betyr enhver praksis som etablerer en slik relasjon mellom elementer slik at elementenes identitet modifiseres (ibid).

---

<sup>10</sup> Et eksempel på et nodalpunkt er ”kroppen” i legevitenskapen. Tegn som ”symptomer”, ”vev” og ”skalpell” får mening i den medisinske diskurs ved at de blir relatert til ”kroppen” på bestemte måter (Winther Jørgensen & Phillips 1999:37).

Oppsummert kan vi si at diskursen alltid kjemper for å fjerne flertydigheter ved å gjøre elementer til momenter gjennom en fiksering av mening. Men denne bestrebelsen lykkes aldri helt fordi de alternative meningene som diskursen fortrenger til det diskursive felt, alltid truer med å destabilisere entydigheten. Alle momenter forblir derfor potensielt flertydige, det vil si at momentene alltid er potensielle elementer. Konkrete artikuleringer reproducerer eller utfordrer de gjeldende diskursene ved å fikse mening på bestemte måter. På grunn av den alltid potensielle flertydigheten er ethvert muntlig eller skriftlig uttrykk til en viss grad en artikulering eller nyskaping. Selv om uttrykket trekker på tidligere fikseringer av mening, altså tidligere diskurser, hvor tegnene har blitt til momenter, så er et uttrykk aldri bare kopi av noe. Ethvert uttrykk bør sees på som en bestrebelse mot å redusere alternative muligheter for mening. Dette skjer fordi man hele tiden setter tegnene i bestemte relasjoner til hverandre og dermed utelukker alternative relasjoner.

Laclau og Mouffes diskursbegrep omfatter som vist ikke bare språk, men alle sosiale fenomener (op.cit:44). Diskursen forsøker å strukturere tegnene som om alle tegn hadde en fast og bestemt mening i en total struktur. Samme logikk gjelder for hele det sosiale området. Man handler som om virkeligheten har en bestemt og entydig struktur, som om samfunnet, de gruppene man tilhører og ens identitet, er objektivt gitt. Men, akkurat som språket aldri er helt gitt, er også samfunn og identitet flytende og foranderlige størrelser som aldri helt kan fikses. Analysens formål er derfor ikke å avdekke den objektive virkelighet, for eksempel å finne ut hvilke grupper samfunnet ”i virkeligheten” består av, men å undersøke *hvordan* man skaper virkeligheten slik at den blir en objektiv og selvfølgelig omverden. Diskursteoriens utgangspunkt er dermed at man skaper *objektivitet*, og det er denne skapelsesprosessen analysen skal følge.

Her er det mulig å hevde at det finnes en sammenheng mellom fokuset i diskursteorien og oppgavens begreper *rasjonalitet*, *legitimitet* og *institusjonalisering*. Det handler om hvordan vi skaper, og har behov for å skape, *mening*, noe som igjen kan sees på som behov for å institusjonalisere en bestemt meningsverden i en organisasjon. Objektivitetsbegrepet i diskursteorien ligger nært opptil min forståelse av institusjonaliseringsbegrepet. Det kan også kanskje betraktes som et begrep som dekker både rasjonalitet, legitimitet og institusjonalisering. Skapelsen av objektivitet handler

om å skape rasjonalitet og legitimitet for noe (en diskurs) og dermed prøve å institusjonalisere dette (denne diskursen) fremfor noe annet (andre diskurser).

Denne meningskonstruksjonen beskrives av Laclau og Mouffe med begrepet *politikk* (op.cit:47). Begrepet forstås ikke i snever forstand som for eksempel partipolitikk: det er derimot et vidt begrep som henviser til at vi hele tiden konstituerer det sosiale på bestemte måter, noe som utelukker andre måter. Det å skape samfunnet på en bestemt måte, og dermed utelukke alle de andre måtene, er det diskursteorien forstår med politikkbegrepet. Politikk er derfor ikke bare en overflate som avspeiler en dypereliggende sosial virkelighet, tvert imot er selve den sosiale organisering et resultat av vedvarende politiske prosesser. På et gitt tidspunkt kan det være kamp om bestemte diskurser, og det blir synlig at forskjellige aktører forsøker å fremme forskjellige innretninger av samfunnet. Andre sosiale praksiser kan forekomme så naturlige at man nesten ikke kan se at det finnes alternativer. De diskursene som er så fast etablerte at man glemmer deres kontingens, blir i diskursteorien betegnet som *objektive* (op.cit:48). I denne sammenheng kan man si at det har skjedd en institusjonalisering av diskursen.

Begrepet *makt* knyttes til politikk og objektivitet i diskursteorien (op.cit:49). Makt forstås her som produsent av det sosiale, det vil si den kraft og de prosesser som skaper vår sosiale omverden og gjør den meningsfull for oss. Maktbegrepet i diskursteorien kan oppfattes som en parallell til den beskrivelsen av rasjonalitetsbegrepet som jeg legger til grunn, hvor det sentrale er hvordan den utdanningspolitiske dagsorden argumenterer for å skape rasjonalitet til kvalitetsvurdering. Makt er dermed ikke noe som noen er i besittelse av, men må forstås som produsent av det sosiale. Makt er et sentralt begrep i diskursteorien da det handler om produksjon av det sosiale og skapelse av vår viten, våre identiteter og relasjoner til hverandre som grupper og individer. Maktbegrepet understreker kontingensen i vår sosiale omverden: viten, identiteter og sosiale relasjoner er alle kontingente. Alle disse er på et tidspunkt gitt på én bestemt måte, men de kunne ha vært, og kan bli, annerledes. Derfor kan man si at makt er produktiv, den produserer det sosiale på bestemte måter.

Makt åpner for institusjonalisering (ibid). Laclau og Mouffes understreker at makt ikke er noe man kan tenke på som borte. Man er avhengige av å leve i en sosial orden, og den er alltid konstituert i makt. På samme tid er man ikke avhengige av å leve i én

bestemt orden. Utelukkelsen av andre ordener er også en del av maktens virkning. Makt produserer en beboelig omverden for oss, samtidig som den avskjærer alternative muligheter. Objektivitet blir dermed *avsporet makt*, hvor maktsporene har blitt visket ut, det vil si der man har glemt at omverdenen er politisk organisert.

Diskursteoriens begreper kan også knyttes til *symbol- og verktøyperspektivet*. Laclau og Mouffes oppfatter makt og politikk som to sider av samme sak: *makt* henviser gjerne til produksjonen av objekter som for eksempel "samfunn" og "identitet", mens *politikk* snarere brukes om den alltid tilstedeværende kontingensen i disse objektene (ibid). *Objektivitet* henviser så til den omverdenen vi tar for gitt, hvor vi har "glemt" at verden alltid konstitueres i makt og politikk. Symbolperspektivet, som utdypes i kapittel 2.8, har også en forståelse av omverden som full av symboler (tegn) hvor man hele tiden prøver å gi disse symbolene rasjonalitet (objektivitet) og se dem i en et rasjonelt perspektiv. Her synliggjøres også, som i diskursteorien, en forståelse for makt som noe som handler om å legitimere én spesiell "oppskrift" eller løsning. Symbolene blir oppfattet som rasjonelle verktøy idet de har oppnådd rasjonalitet og legitimitet. Dermed har de en mulighet til å institusjonaliseres i en organisasjon. Oppsummerende kan vi si at både diskursteori og symbol- og verktøyperspektivet handler om å konstruere *mening*.

## 2.8. Symbol- og verktøyperspektivet

Det er et mål i denne oppgaven å skape forståelse for at kvalitetsvurdering er et komplekst fenomen. Med dette vil jeg prøve å bidra til å utvide forståelsen av kvalitetsvurdering, og fenomenets inntog på den utdanningspolitiske dagsorden.

Med utgangspunkt i et symbolperspektiv kan man muligens hevde at dårlig kvalitet er et skapt problem snarere enn et organisasjonsinternt problem. Dette betyr at det kan være et problem som skapes utenfor selve organisasjonen, men som likevel føles på "kroppen" internt. Med dette perspektivet vil jeg vise til de tendenser i samfunnet som er med på å påvirke den utdanningspolitiske dagsorden, og som dermed er med på å styre hvilke valg som blir fattet. *Kvalitetsvurdering* bør i denne sammenheng betraktes

som et fenomen som kan være med på å gi et bilde av hvilken retning utdanningspolitikken beveger seg i. *Hvorfor kvalitetsvurdering har utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden* er derfor et sentralt spørsmål, i tillegg til mitt andre spørsmål i problemstillingen: *Hvilken argumentasjon legges til grunn for kvalitetsvurdering på den utdanningspolitiske dagsorden?*

Her kan man hevde at forholdet mellom verktøy- og symbolperspektivet står sentralt. Dette er to forskjellige perspektiver som kan være med å forstå organisasjoners atferd og hvordan visse ”oppskrifter” på en god organisasjon utvikles og forstås. Disse perspektivene kan sees på som perspektiver som kan hjelpe diskusjonen på vei til å forstå hvorfor kvalitetsvurdering har blitt et sentralt element på den utdanningspolitiske dagsorden. I tilknytning til dette er det rimelig å anta at man kan si noe om hvordan kvalitetsvurdering institusjonaliseres, eller forsøkes institusjonalisert, som ”det eneste riktige å gjøre” for å hankses med utdanningspolitiske utfordringer og oppgaver.

Røvik mener at verktøyperspektivet har sitt opphav i den rasjonelle-instrumentelle tradisjon i organisasjonsteorien (Røvik 1998). Tradisjonen kan tolkes som en fellesbetegnelse som knytter sammen flere vitenskapelige retninger og forskere fra ulike perioder på 1900-tallet. Utviklingen fra scientific management, gjennom Webers klassiske byråkratilære, og til statsvitenskap og fokusering på effektivitet og kvalitet i produksjon, kan skissere et bilde av utviklingen frem til verktøyperspektivet. Felles for disse retningene er at de ser organisasjoner kun som redskaper for effektiv frembringelse av vedtak, varer, tiltak og tjenester. Dermed kan man si at organisasjoners legitimitet i dette perspektivet er knyttet til hvor effektive de klarer å være. I verktøyperspektivet kan det viktigste redskapet for effektivitet oppfattes å være *planmessig organisasjonsutforming*, det gjelder å strebe etter å forutse konsekvenser av å velge den ene eller andre løsningen. Her er en organisasjonsoppskrift kun et redskap, det vil si at den sees på som ”et av flere mer eller mindre velprøvde og av erfaring finslippte verktøy som ledelsen kan bruke i det kontinuerlige arbeidet med å gjøre organisasjonen mer effektiv.” (op.cit:32).

Fordi verktøyperspektivet kan tolkes som preget av utviklingsoptimisme og en sterk rasjonalistisk orientering, peker det seg ut som modernistisk inspirert. Her kan man finne en sterk tro på forandring og forbedring ved hjelp av organisasjoner og aktører

som er kunnskapsrike, og som har både læringsvilje og læringsevne. Dette kan tyde på at opplysningstidens fornuftbegrep og modernitetens tro på fremskritt lever videre i verktøyperspektivet.

På slutten av 1900-tallet ble verktøyperspektivet utfordret av bidrag innen nyinstitusjonell teori (op.cit:35). Dahler-Larsen forklarer at innenfor nyinstitusjonell teori kan ikke livet i organisasjoner reduseres til, eller forklares ut fra, organisasjoners formelle mål eller gruppers snevre egeninteresser (Dahler-Larsen 2001a:46). I tillegg kan man se utfordringer fra etnometodologisk inspirerte organisasjonsforskere og fra den europeiske konstruktivistiske tradisjon. Samlet kaller Røvik denne retningen for symbolperspektivet (Røvik 1998:36). Det grunnleggende i dette perspektivet er i følge Røvik at formelle organisasjoner befinner seg i institusjonelle omgivelser der de hele tiden møter sosialt skapte normer og konvensjoner for hvordan den enkelte organisasjon til enhver tid bør være utformet. Han henviser til Berger og Luckmann for å vise at institusjonelle omgivelser ikke er naturgitte, objektive realiteter, men heller samfunnsskapte fenomener (ibid). Disse egentlig sosialt konstruerte normene for hva som er god organisering, oppfattes gjerne som ytre, objektive og gitte betingelser. Denne transformasjonen blir mulig gjennom institusjonalisering (ibid). Videre bruker Røvik begrepet institusjonalisering for å vise de prosessene som fører til at en organisasjonsoppskrift gradvis fremstår som et regelliknende faktum, det vil si noe som tas for gitt som en effektiv og tidsriktig måte å organisere på. Dermed følger det at organisasjoners legitimitet avhenger av blant annet hvorvidt man er i stand til å bygge inn og reflektere de til enhver tid gjeldende institusjonaliserte forestillinger i omgivelsene. Røvik hevder at disse institusjonelle omgivelsene ikke kan sees på som stabile. Normer for hva som er en moderne organisasjon endrer seg, og organisasjonene opplever derfor et vedvarende høyt endringstrykk. Dette endringstrykket kan komme til uttrykk i ulike forslag til stadig nye reformer (ibid).

I dette perspektivet fungerer de institusjonaliserte organisasjonsoppskriftene som meningsbærende symboler. De gis et meningsinnhold som strekker seg lengre enn bare å være verktøy for effektiv problemløsning. Symbolene er *rasjonaliserte* ved at det legges vekt på å definere og fremstille dem som finslipte redskaper for effektivisering og modernisering. Røvik skriver:

”Påstanden er at vår tids populære organisasjonsoppskrifter har fått sin legitimitet og spredningskraft nettopp ved at de er blitt symboler på grunnleggende rasjonalistiske verdier i det moderne samfunnet, som fornuft effektivitet, styring, demokrati og vitenskapelighet, og slik sett er blitt assosiert med det kontinuerlige fremskrittet, dvs. bevegelsen fremover og oppover mot noe stadig bedre.” (ibid).

Her knytter Røvik symbolperspektivet til verktøyperspektivet og den rasjonelle-instrumentelle tradisjon på en måte som gjør det vanskelig å snakke om enten-eller.

Men likevel er det legitimt å skille klart mellom perspektivene. Røvik viser dette i det han kaller problembetinget søking (op.cit:39). I verktøyperspektivet betyr dette at det *først* oppstår et lokalt ”ekte” problem internt i organisasjonen, og at man *deretter* starter letingen etter oppskrifter som kan løse det. I symbolperspektivet arter dette seg annerledes: *først* blir noen i organisasjonen oppmerksom på en populær oppskrift, *deretter* opplever man at man har et lokalt ”problem” må løses. Dette forklarer hvorfor mange organisasjoner omtrent samtidig opplever at de har de samme typer problemer, som for eksempel problemer med dårlig kvalitet i interne funksjoner, og at de forsøker å løse dem med omtrent den samme type oppskrift.

Røvik bemerker at det fortsatt er slik at man kan snakke om et rasjonalistisk-instrumentelt hegemoni. De to perspektivene er på ingen måte likeverdige når det gjelder status og aksept. Den rasjonalistiske tilnærmingen etablerer legitime roller, tilbyr et eget rasjonalisert språk, definerer arenaer og setter rammer for de fleste av samtidens diskurser om organisasjons- og ledelsesspørsmål. Fra verktøyperspektivet oppfattes som oftest de to perspektivene som direkte motpoler. Forholdet mellom dem fortolkes gjerne normativt som et skille mellom den riktige og den gale orientering, mellom det man ønsker seg (gode, pålitelige verktøy) og det man ønsker å unngå (moter og uekte varer). Argumenter som dette bygger på den rasjonalistiske forestillingen om at det er mulig, og nødvendig, å skille skarpt mellom symboler og ekte verktøy, mellom det ytre og estetiske og de indre og egentlige organisatoriske behov, mellom motehandlere og ekte organisasjonsdoktorer og mellom retorikk og realiteter (op.cit:41).



### 3. BEGREPSDRØFTING

I dette kapittelet vil jeg presentere og drøfte begrepene ”kvalitet” og ”evaluering”. Man kan forstå kvalitetsvurdering som et fenomen sammensatt av nettopp disse to. Ved å gå dypere inn i begrepene, vil jeg forsøke å belyse den kompleksiteten som kan sies å ligge i dem. Man kan argumentere for at kvalitets- og evalueringsbegrepet har forskjellige historier. Evalueringsbegrepet kan sies å ha blitt utviklet gjennom forskning og teoridannelse, mens kvalitetsbegrepet neppe kan oppfattes som et begrep utviklet gjennom forskning. I oppgaven er det viktig å drøfte begge begrepene: Kvalitet forankres mot den utdanningspolitiske dagsorden og evaluering forankres mot vitenskapsteorien. Videre kan man si at begrepene sammenfattes i en utdanningspolitisk diskursiv kamp.

I tillegg til denne drøftingen vil jeg til slutt presentere og redegjøre for hvordan begrepene rasjonalitet, legitimitet og institusjonalisering blir brukt i denne oppgaven. Man bør se alle disse begrepene som sentrale for å belyse oppgavens problemstilling.

#### 3.1. Hva er *kvalitet*?

Kvalitetsbegrepet kan gjerne oppfattes som problematisk å definere. Spørsmålet kan sees på som en filosofisk oppgave å besvare, og man kan i historien finne mange forsøk på å gi svar på *hva* kvalitet er. Begrepet relateres ofte til det som er bra og verdifullt, dermed kan man kanskje si at det å gi et fullgodt svar på spørsmålet, krever en gjennomgang av alt som er skrevet om *det gode* og *det verdifulle*. Dette er verken mulig eller ønskelig i denne sammenheng, men det kan likevel være mulig å foreta noen refleksjoner omkring begrepets mening og innhold.

Granheim gjør et forsøk på å definere, eller avgrense, kvalitetsbegrepet (Granheim 1990b:13). Hun skriver at ”kvalitet” opprinnelig stammer fra det latinske ordet *qualitas* og kan oversettes med ”hvordan sammensatt”, i betydningen ”som objektivt virkelig er”. Begrepet har både et objektivt innhold som knyttes til et bestemt produkt eller en tjeneste, og et subjektivt innhold som knyttes til brukernes oppfatning av produktet eller

tjenesten, dets evne til å tilfredsstille hans/hennes spesielle ønsker, forventninger og behov. Videre skriver hun: "På undervisningssektoren, der 'tjenesten' som tilbys er beskrevet ved et utall generelle og spesifikke mål i lover og læreplaner, og der flere av målene er motstridende, blir derfor det å definere kvalitet ingen enkel oppgave." (ibid).

Hvis man ser på *kvalitet* som et språk tegn, er det mulig å analysere tegnet ut fra tre dimensjoner. Man kan se på *syntaksen*, det vil si at man studerer relasjonen mellom tegn. *Semantikken* viser til en studie av relasjonen mellom tegn og deres referenter, det vil si at man leter etter hva ordet betyr når det sees i sammenheng med objekter, handlinger eller egenskaper. Den tredje dimensjonen omhandler en *pragmatisk* studie av ordet. Da ser man på hvordan språk tegn brukes i et kommunikasjonsforhold, det vil si hvordan det brukes av en sender og/eller forstås av en mottaker. Det mest interessante her er de to sistnevnte dimensjonene av "kvalitet" som språk tegn: Hva betyr kvalitet når det sees i forhold til mulige referenter, og hvordan blir kvalitet brukt i en kommunikasjonshandling?

Man kan finne utallige meninger som kan festes til kvalitetsbegrepet. Selv om det i dag ser ut til å være enighet om at mening er mer eller mindre subjektiv og avhengig av kontekst, kan man se forøk på å finne en felles forståelse for hvordan ord blir brukt, hvordan de settes sammen med ting, handlinger eller egenskaper (Thune 2000:124). I konkrete tegn som en sykkel eller et brød, er det lettere å finne en felles forståelse enn i mer abstrakte tegn som skjønnhet, sannhet eller kvalitet. For å fikse tegnenes mening må man definere dem (ibid). I ordbøker kan man lete etter forsøk på å meningsfikse tegn ved å peke på like eller overlappende begrepet og deres antonymer, i tillegg til den konteksten som tegnet vanligvis viser seg i. Thune skriver at disse sammenliknende begrepene kan kalles *tegnets semantiske nettverk*, et nettverk som inneholder mulige referenter i form av synonymer og konnotasjoner (ibid).

I kvalitetsbegrepets semantiske nettverk kan man hevde at "kvalitet" har to aspekter. Man kan for det første betrakte kvalitet i forhold til spesifikke oppgaver eller karakteristikk til et fenomen. I det andre aspektet kan man hevde at kvalitet oppfattes som en verdi, det vil si som det man mener er godt eller verdifullt. Thune poengterer at kvalitet også kan vise til egenarten av et fenomen eller hvordan et fenomen er bedre en noe annet. Dette kan tolkes dit hen at det ligger et komparativt element i forståelsen av

kvalitet, noe som betyr at kvalitet peker på at noe er bedre enn noe annet. I denne sammenheng kan man si at kvalitet er et begrep som betegner noe positivt. Hva man velger å vektlegge når man beskriver kvalitet, påvirker forståelsen av hvordan kvalitet brukes i for eksempel *kvalitetsvurdering* på den utdanningspolitiske dagsorden.

Diskusjonen omkring kvalitetsbegrepet kan spores langt tilbake i historien, helt til antikkens filosofer. Ideen om kvalitet som noe verdifullt bringer frem spørsmålet om hva som er godt, eller hva som er essensen i godhet. Man kan skille mellom platonske og aristoteliske svar på dette spørsmålet, noe som skisserer et grunnleggende skille mellom kvalitet som et idealistisk fenomen og kvalitet som et pragmatisk fenomen. Platon mente kvalitet var noe idealistisk og utopisk, det var idealet man skulle strebe etter. I denne forståelsen blir kvalitet noe som gjennomsyrrer objektet, og det kan ikke måles fordi idealet aldri er tilgjengelig (Pirsig 1997<sup>11</sup>). Aristoteles forstod kvalitet som noe pragmatisk og funksjonelt, eller som noe som måtte sees i forhold til et objekt. Med en slik forståelse innehar en handling eller et menneske kvalitet hvis det utfører sin funksjon godt. I aristotelisk tradisjon kan kvalitet måles (ibid).

Skillet mellom disse forståelsene av kvalitet kan fremheves ytterligere ved å vise til uenigheten om kvalitet kan være *a priori* definert, altså om kvalitet er noe som tilhører ideen om objektet eller om kvalitet bestemmes i relasjon til objektet det er snakk om. I den sistnevnte betraktningen ser man på kvalitet som en funksjon i forhold til en observasjon eller handling, og følgelig som en funksjon som den som observerer innehar. Denne distinksjonen bør man merke seg. Det å gi mening til kvalitet kan i lys av dette enten sees på som en direkte funksjon som objektet innehar, og som derfor er bundet til objektet i seg selv, eller som en funksjon som viser til den prosessen hvor en person vurderer et objekt. I denne siste betydningen kan man si at kvalitet ligger utenfor objektet. Man kan derfor hevde at denne *pragmatiske* forståelsen av kvalitet baserer seg på en subjektiv forestilling av kvalitet, mens den *idealistiske* forståelsen ser kvalitet som noe absolutt – *idealet*.

Kvalitetsbegrepet kan sies å ha blitt et sentralt begrep i organisasjonsteorien i etterkrigstiden (Røvik 1998). Kvalitet viste seg som del av *the scientific management*

---

<sup>11</sup> En norsk utgave som er oversatt fra engelsk. Originalen ble utgitt første gang i 1974.

*movement* i japansk industri på 1950-tallet, noe som har blitt kalt *Industrial Quality Movement* (Thune 2000:127). Her ble kvalitet operasjonalisert som ”conformance to specifications” og sett på som en del av den helhetlige produksjonsprosessen som skulle forhindre ustabilitet og svinn. Man kan hevde at denne tilnærmingen til kvalitet har et teoretisk fundament i statistiske analyser av årsaker til avvik. Disse ideene har vist seg i flere organisasjonsoppskrifter, blant annet i Total Quality Control (TQC) og Total Quality Management (TQM). Alle disse konseptene har det til felles at de relaterer kvalitet til markedssuksess. Her kan man oppfatte to aspekter som fremtredende: fokus på produksjonsprosessen og fokus på krav fra kunden. Men i tillegg til fokuset på funksjonalitet, kan man hevde at den viktigste kvalitetsmålingen sees i forhold til om kundene til slutt kjøper produktet eller ikke. Derfor kan man si at denne forståelsen av kvalitet kan forstås som en del av den pragmatiske dimensjonen (op.cit:128).

Når kvalitet brukes i et produksjons- og markedsperspektiv kan man se begrepet som en markedsstrategi ved at produkter merkes med ”kvalitet” eller ”kvalitetssikret”<sup>12</sup>. Thune viser til Brennans kommentarer om hvordan kvalitet brukes i samfunnet i dag: ”As a prefix, it attaches itself to everything: from automobiles to baked beans, from plumbing to lager. The adjectival use of the term evokes the marked place, the advertising hoarding, the ‘ad’ in Yellow Pages. It is part of the non-too-subtle ‘hard sell’.” (ibid).

Med disse dimensjonene som et rammeverk er det kanskje mulig å si noe om hvordan kvalitet blir brukt på den utdanningspolitiske dagsorden i dag.

### 3.1.1. Kvalitetsbegrepet i bruk på den utdanningspolitiske dagsorden

Man kan muligens hevde at det i den åpne politiske debatten om utdanning er vanskelig å gripe fatt i den pragmatiske dimensjonen i kvalitet. Det kan virke som om kvalitet er et begrep som opptrer kun i idealistisk form, og at det er umulig å bruke begrepet pragmatisk på utdanningsfeltet.

---

<sup>12</sup> Fra engelsk ”quality” og ”quality approved”.

På den utdanningspolitiske dagsorden har det i de siste årene pågått en diskusjon om hvordan man skal innholdsdefinere kvalitet i utdanningen. Arbeidet med å gi utdanningskvalitet et innhold har vært en av kvalitetsutvalgets<sup>13</sup> hovedoppgaver (NOU 2003:16). I NOU 2002:10 definerer kvalitetsutvalget kvalitet gjennom tre områder som igjen kan spesifiseres i kvalitetskriterier. Dette kan tolkes dit hen at man vil prøve å meningsfiksere kvalitet ved å definere kvalitetsstandarder som deles inn i tre områder. For det første vises det til de ytre rammene for undervisningen i form av ressurser, som for eksempel bevilgninger, antall lærere per elev, lokaler og utstyr. Denne kvalitetstypen omtales som *ressurskvalitet*. Det andre kvalitetsområdet vektlegger selve læringsprosessen, og omfatter blant annet innholdet i undervisningen, hvordan undervisningen er lagt opp og hvilke metoder som benyttes. Denne kvalitetstypen kalles *prosesskvalitet*. Det siste området omtales som *resultatkvalitet*. Denne kvalitetstypen skal si noe om elevens eller studentens utbytte av utdanningen, ikke bare i form av karakterer, men i form av et helhetlig læringsutbytte (NOU 2002:10).

Denne oppdelingen i tre kvalitetsområder og tilhørende kriterier kan tolkes som en del av den pragmatiske dimensjonen av kvalitetsbegrepet.

### 3.2. Evalueringsbegrepet

Evalueringsbegrepet kan sees på som et begrep som fordrer en form for verdisetting; man må ta stilling til verdier, normer eller kriterier når man skal vurdere verdien av en evaluand. Her kan man argumentere for at det finnes en kopling til kvalitetsbegrepet. Man kan hevde at det finnes en ubrytelig sammenheng mellom kvalitet på den ene siden og spørsmålet om evaluering på den andre. Når kvalitet blir omtalt, følger vanligvis en eller annen form for evaluering med. Følgelig kan man hevde at disse to henger tett sammen; det ene kan sees på som en forutsetning for, og kanskje garanti for, det andre.

Dahler-Larsen mener det er behov for en mer sofistikert forståelse av evalueringsfenomenet<sup>14</sup>. Tolkning av ulike definisjoner er en måte man kan angripe en

---

<sup>13</sup> Kvalitetsutvalget ble oppnevnt 2001. Mandatet til utvalget var blant annet å komme med forslag til tiltak for å forbedre kvaliteten i norsk grunnsopplæring.

<sup>14</sup> Se kapittel 1.1.

slik utfordring, selv om dette bare kan sies å være en start. I dette kapitlet vil jeg presentere fem definisjoner av begrepet ”evaluering” og kommentere dem kort. Presentasjonen kan kritiseres for måten *educational evaluation*, *evaluation reseach* og *evaluation/evaluering* ser ut til å sidestilles. Hensikten er ikke å sette likhetstegn mellom disse, men heller å gi et bilde av at evaluering brukes på forskjellige måter og på ulike områder. I tillegg kan en slik fremstilling muligens være med på å vise noe av den kompleksiteten som ligger i ”evaluering”.

Evalueringsens plass på den utdanningspolitiske dagsorden vil bli presentert og drøftet utover i oppgaven.

Den første definisjonen er hentet fra Gall, Borg og Gall og er en definisjon som stammer fra utdanningsfeltet: ”Educational evaluation is the process of making judgements about the merit, value, or worth of educational programs.” (Gall m.fl. 1996:680). Denne definisjonen kan tolkes dit hen at den er for vid til å bruke som en arbeidsdefinisjon.

Den neste definisjonen er hentet fra Judd, Smith og Kidder:

”Evaluation research is best thought as one type of applied research in the sense that it is designed to answer practical real-world problems about the effect of some policy or program. Although applied research more broadly may serve to define a social program or explore alternative policies or programs that might be implemented to solve some problem, evaluation research has as its primary goal the description and evaluation of some existing social policy or program.” (Judd m.fl. 1991:322).

Etter min oppfatning overfokuserer denne definisjonen på effekt. Den kan tolkes som en definisjon som snevrer inn forståelsen på en negativ måte, det vil si at vi mister en del av det evaluering kan være uten om å gjelde effektmålinger.

Weiss definerer evaluering på følgende måte: “Evaluation is the systematic assessment of the operation and/or the outcomes of a program or policy, compared to a set of explicit or implicit standards, as a means of contributing to the improvement of the program or policy.” (Weiss 1998:4). Her kommer det frem at vi må ha kriterier/standarder for å vurdere om evalueringsresultatene er gode eller dårlige. Likevel mener jeg at det ikke er en fullgod definisjon da kontrollelementet

undervurderes. Definisjonen kan sies å underkommunisere at enhver evaluering har et kontrollelement i seg, i tillegg til et utviklingsselement.

Scriven definerer evaluering som "the determination of the worth or value of something – in this case, of educational and social programs, policies and personnel – judged according to appropriate criteria explicated and justified." (Scriven 1991b). Dette er en av de mest brukte definisjonene i faglitteraturen. Den er, etter min mening, en god definisjon fordi det kommer tydelig frem at evaluering har ulike dimensjoner i seg. Hovedvekten i definisjonen ligger på verdiaspektet i evaluering, det handler om å bestemme om det en står overfor er godt eller dårlig, om det er slik det burde være, eller ikke.

Den siste definisjonen som presenteres er en definisjon av Vedung. Dette er en velkjent og mye brukt definisjon i de skandinaviske landene. "Utvärdering" er en "noggrann efterhandsbedömning av utfall, slutprestationer eller förvaltning i offentlig verksamhet, vilken avses spela en roll i praktiska beslutssituationer." (Vedung 1998:20). Dahler-Larsen og Krogstrup bruker denne definisjonen som en arbeidsdefinisjon. De mener definisjonen kan diskuteres, men den viser vesentlige dimensjoner som finnes i enhver evalueringsproblematikk (Dahler-Larsen & Krogstrup 2001b). Man kan kritisere definisjonen for blant annet å bruke begrepet *efterhandsbedömning* (tilbakeskuende), noe som kan bety at fremadskuende teknikker ikke er med. I tillegg kan den kritiseres for å innsnevre evaluering til kun å gjelde offentlig politikk. Men definisjonen er likevel videre enn for eksempel en del amerikanske definisjoner som opererer med en fastsatt kjøreplan bestående av: etablering av kriterier for innsatsen, konstruksjon av standarder, måling og sammenlikning med standarder.

Vedungs definisjon viser til at evaluering er systematisk. Det betyr at evaluering er noe annet enn bare synspunkter. Det ligger en viten bak evalueringen, en viten som er systematisk produsert ved hjelp av en eller annen metode (bringer evaluering i familie med vitenskap). For det andre er evaluering en vurdering som logisk forutsetter noen verdier. For det tredje viser definisjonen til en tilsiktet anvendelse i praktiske handlingssituasjoner som politiske beslutninger eller fagprofesjonell praksis. For det fjerde viser definisjonen til en evaluand, det vil si et nærmere bestemt aspekt av en offentlig politikk.

Shadish, Cook og Levitons betraktninger omkring temaet kan sees i sammenheng her. De mener også at en velargumentert evaluering i praksis er et resultat av overveielser av nettopp *viten, verdier, anvendelsen og evaluanden* (Shadish m.fl. 1991). Herved sies det ikke at all evalueringspraksis er velargumentert. Oppgaven med å bringe evalueringspraksis i logisk overensstemmelse med alle gode argumenter om henholdsvis *viten, verdier, anvendelse og evaluanden*, er en utfordring man ikke bør se lett på. Dette skyldes ikke minst at det i dag er sterke utviklinger på spill i hver av de fire dimensjonene (Dahler-Larsen & Krogstrup 2001b):

I *vitensdimensjonen* ser man at evaluering følger den alminnelige samfunnsvitenskapelige utvikling hvor klassiske positivistiske idealer både erstattes og suppleres med hermeneutiske og fenomenologiske idealer og postmoderne tankefigurer. Dette medfører en voldsom utvikling i de anvendte metodene. Samtidig prøver man å fastholde det vitenskapsliknende elementet i evaluering, noe som betyr at man kan stille validitetskrav til evalueringsviten. I dag er det store diskusjoner om hvilke vitenskapskrav som kan stilles til evalueringer, og i hvilket omfang evaluering overhodet er i familie med vitenskap.

*Verdidimensjonen* har også eksplodert. Prinsipielt kan evaluatoren enten legge relevante involverende parter verdier til grunn eller fremføre sine egne. I evalueringens barndom henviste man typisk til den parlamentariske styringskjede, representert ved ansvarlige politikere, når man skulle finne et relevant verdigrunnlag for en evaluering. Dette verdigrunnlaget har ikke lenger monopol. Det aktualiserer debatten om hvilke involverte parter som legitimt kan velges ut som leverandører til verdigrunnlaget for evalueringen. Herved skjerpes også kravene til evaluator med hensyn til å begrunne hvilke verdier som velges. Disse kravene blir naturligvis ikke mindre når evaluator bevisst eller ubevisst trekker med seg sine egne verdier inn i evalueringen.

Utviklingen i *anvendelsesdimensjonen* kan også karakteriseres som dramatisk. Man er nå blitt oppmerksom på at ikke bare evalueringsresultatene, men også evalueringsprosessen kan finne ”anvendelse”. Antallet anvendelsesformer er tilsvarende utvidet. Oppmerksomheten er nå på flere områder; de samfunnsmessige, politiske, institusjonelle og organisatoriske sammenhengene som evalueringen finner sted i.



*Evaluanden* kan i dag være mange forskjellige ting. Vi har utviklet oss langt bort fra evalueringsbegrepets fødsel i USA hvor evaluanden typisk var et eksperimentelt program hvis fortsettelse man skulle vurdere. I dag kan man evaluere policies, programmer og programkomponenter. Man kan evaluere strukturer, prosesser og resultater. Man kan evaluere personer, organisasjoner, samarbeidsformer og styreformer. Man kan undersøke ikke bare om bestemte aktiviteter fører til bestemte resultater, men også problematisere om det er det riktige som gjøres. Den tilsynelatende samme innsats kan betraktes på mange forskjellige måter, alt etter hvordan evaluanden blir definert.

### 3.3. Rasjonalitet, legitimitet og institusjonalisering

Rasjonalitet, legitimitet og institusjonalisering bør sees på som sentrale begreper i forhold til problemstillingen i denne oppgaven. Institusjonalisering kan sies å være et sentralt begrep fordi et tiltaks virke og varighet avhenger av om det blir institusjonalisert i systemet. Rasjonalitet og legitimitet kan oppfattes som sentrale i forhold til om et tiltak vil bli institusjonalisert. Det er rimelig å anta at institusjonalisering avhenger av om man klarer å skape en rasjonalitet omkring tiltaket, og om man oppnår legitimitet.

De tre begrepene kan sies å ha betydningsmessig slektskap da alle tre forstås i en sammenheng som har med forståelse og aksept å gjøre. I tillegg er det mulig å si at det er menneskelig å tenke i disse begrepene da de skaper trygghet og tilhørighet. Man kan hevde at begrepene handler om meningskonstruksjon, objektivisering og institusjonalisering.

Rasjonalitetsbegrepet kan ha forskjellige betydninger. I denne oppgaven forstås ikke rasjonalitet som rasjonalisering, det vil si ”å få utført så mye som mulig ved hjelp av de midlene som er til rådighet”, men som å gi noe rasjonalitet ved å tillegge det rasjonell mening. De rasjonalistiske verdiene som fortsatt gjør seg gjeldende i det moderne samfunnet, kan sies å være viktig for at oppskrifter og løsningsforslag velges og brukes.

Løsningene og organisasjonsoppskriftene kan sees på som symboler på fornuft, effektivitet, styring, demokrati og vitenskapelighet, og kan videre assosieres med det kontinuerlige fremskrittet og forstås som det eneste rasjonelle å gjøre. Symbolene bør forstås som rasjonaliserte symboler som strever for å bli definert og fremstilt som finslippte redskaper for effektivisering og modernisering (Røvik 1998:36). I forbindelse med dette er det mulig å si at løsningene tildeles rasjonalitet, da de blir assosiert med fremskritt og forbedring.

Røvik hevder at organisasjoners legitimitet avhenger blant annet av hvorvidt man er i stand til å bygge inn og reflektere over de til enhver tid gjeldende institusjonaliserte forestillingene i omgivelsene (ibid). Denne tankegangen kan man finne igjen i symbolperspektivet hvor populære ideer og oppskrifter på god organisering er noe mer enn kun effektivitetsredskaper. Ideenes og oppskriftenes spredning avhenger av om de viser seg som meningsbærende symboler på utbredte normer og verdier som råder i samfunnet, slik som rasjonalitet, effektivitet, fornyelse, utvikling, demokrati, individualisme og rettferdighet (op.cit:48). For å belyse dette kan man referere til Talcott Parsons. Han hevdet i en artikkel i 1956 at moderne organisasjoner ikke ville overleve ved bare å være effektive i teknisk-økonomisk forstand (ibid). Parsons så et behov for legitimitet i tillegg. Videre hevdet han at for å fremstå som legitime måtte organisasjonene inkorporere og utad signalisere en tilslutning til verdier og normer i sitt handlingsmiljø. Denne forståelsen kan oppfattes som *kjernen* i den nyinstitusjonelle skole i organisasjonsteorien, en forståelse hvor man betrakter organisasjonens omgivelser fra et klassisk *innenfra og ut-perspektiv*. I denne oppgaven, med referanse til symbolperspektivet og Røvik, argumenteres det for at det samme resonnementet gjelder fra et *utenfra og inn-perspektiv*, det vil si der man betrakter organisasjonene fra omgivelsene. Røvik mener vi får et tydeligere bilde av institusjonelle omgivelser ved å bruke det han kaller ”en omvendt Parsons”. Dette innebærer at man legger til grunn et utenfra og inn-perspektiv og studerer innholdet i de ideer og oppskrifter som har den egenskapen at de aksepteres og spres svært lett og raskt i organisasjonssamfunnet (op.cit:49).

I organisasjonsteorien brukes institusjonaliseringsbegrepet for å vise hvilke prosesser som fører til at noe i en organisasjon gradvis fremstår som et regelliknende faktum, det vil si som noe som tas for gitt som en effektiv og riktig måte å organisere på (op.cit:19).

Man kan muligens i forlengelsen av dette hevde at institusjonalisering avhenger av om rasjonalitets- og legitimitetsprosjektet lykkes både nedover i organisasjonen og ut i det offentlige.

Det er rimelig å anta at disse tre begrepene har betydning for utviklingen frem mot et kvalitetsvurderingssystem på den utdanningspolitiske dagsorden. Man må argumentere slik at det skapes en rasjonalitet til kvalitetsvurdering, og videre argumentere slik at den oppnår legitimitet. Først da kan man snakke om at systemet er på vei til å bli institusjonalisert.

## 4. UTVIKLINGSLINJER

I kapittel tre drøftet jeg kvalitets- og evalueringsbegrepet, i tillegg til begrepene rasjonalitet, legitimitet og institusjonalisering. I dette kapitlet vil jeg vise til historien for å belyse evalueringsfenomenets utvikling og se på hvordan det har blitt vektlagt til forskjellige tider. En historisk fremstilling er nødvendig da denne kan sies å danne basis for diskusjonen om hvilke forhold som influerer kvalitetsvurderingsfokuset på den utdanningspolitiske dagsordenen. Målet med denne fremstillingen er å gi et bilde av hvilke diskursive kamper som har utspilt seg, og hvilke diskurser som kan sies å ha lyktes i den hegemoniske intervensjonen<sup>15</sup> opp gjennom historien. Det historiske bildet fremstilles i både et nasjonalt og internasjonalt perspektiv. Man kan hevde at norsk utdanning har blitt inspirert fra ulike steder til ulike tider når det gjelder evaluering. Norge hentet sin inspirasjon fra Sverige og Storbritannias *School Councils* på 1960- og 1970-tallet. Spesielt kan man her nevne Stenhouse, og læreren som forsker, og lokalt utviklingsarbeid. Det var skolevurdering på de lokale nivå som stod i sentrum. På slutten av 1980-tallet og begynnelsen av 1990-tallet kan man oppfatte en endring i den skolepolitiske tankgangen til land som Norge lenge hadde sett på som inspirasjonskilder. Telhaug og Mediås nevner spesielt *thatcherismens* skolereformer fra 1988<sup>16</sup> og årene fremover (Telhaug & Mediås 2003:291). Disse reformene i Storbritannia kan sees på som en innsnevring av de lokale myndighetenes og den enkelte skoles frihet, da sentraliserte vedtak om en bindende, nasjonal læreplan, prestasjonsmål for de enkelte fagene, vurdering og inspeksjon ble gjeldende.

Det er rimelig å hevde at denne utviklingen også påvirket, om noe senere, den utdanningspolitiske dagsorden i Norge. I tillegg kan man argumentere for at endringer i samfunnet også har vært og er med på å utforme den utdanningspolitiske dagsorden<sup>17</sup>. Når det gjelder evalueringshistorien er det naturlig å bevege seg vestover, til USA, for å belyse hvordan evalueringsfeltet utviklet seg som et forskningsfelt. Når jeg nå vektlegger den amerikanske historien er ikke det fordi den utdanningspolitiske

---

<sup>15</sup> Den hegemoniske intervensjonen lykkes hvis en diskurs står igjen og dominerer i det feltet der det tidligere var konflikt (antagonismen er dermed oppløst). En fastlåsning av mening har skjedd, det vil si at meningskonstruksjonen er objektivert/institusjonalisert (Winther Jørgensen & Phillips 1999).

<sup>16</sup> The Educational Reform Act av 1988 inneholdt mer enn 200 paragrafer. Disse gav utdanningsministeren en ny makt som han kunne regulere nesten ethvert aspekt ved utdanningssystemet med (Telhaug & Mediås 2003:296).

<sup>17</sup> Dette vil bli drøftet i kapittel fem.

dagsorden hentet sin inspirasjon direkte herfra, men for å vise hvilken intellektuell utvikling feltet hadde, og hvordan dette kan sies å ha influert norsk tenkning indirekte. Fra norsk hold har man kanskje ikke sett på USA som den store inspiratoren, men heller holdt seg til de europeiske landene. Man kan imidlertid hevde at det er mulig å se relativt klare tegn på at den amerikanske utdannings- og evalueringsteorien influerte norsk utdanningspolitikk i perioder. Dette kan man se blant annet i artikkelsamlingen *Utdanningskvalitet – styrbar eller ustyrlig?*, som er et resultat av arbeidet etter OECDs tilstandsrapport om norsk utdanningspolitikk<sup>18</sup> (Granheim m.fl. 1990).

Artikkelsamlingen er en dokumentasjon av to prosjektfaser i arbeidet etter OECD-rapportens ankomst<sup>19</sup>. For det første ønsket man å danne seg et bilde av internasjonale kunnskaper og erfaringer om hvordan erindringer i den politiske styringen av skolesektoren hadde påvirket synet på evaluering, og hvilke konsekvenser dette kunne få for relasjonene mellom nivåene i sektoren. Dette resulterte i at man inviterte internasjonalt anerkjente eksperter til et seminar i Oslo i mai 1989. Utgangspunktet for diskusjonene var en artikkel Lundgren hadde skrevet om utfordringene det desentraliserte norske utdanningssystemet sto overfor ved å innføre evaluering. I den andre fasen ble de viktigste erfaringene og kunnskapene som fantes i Norge på dette feltet analysert og dokumentert (Granheim 1990a:9).

Da det var flere amerikanske evalueringsteoretikere som deltok på seminaret, er det rimelig å anta at dette prosjektet, spesielt med den første fasen, i en viss grad ble påvirket av amerikanske ideer.

Den amerikanske evalueringshistorien er også relevant fordi den kan sies å ha betydning for en utvikling man nå kan øyne i Norge, en utvikling som ser ut til å gå i retning av standardisering av kvalitetsindikatorer og resultatstyring.

I tillegg til en historisk fremstilling, vil jeg redegjøre for, og diskutere, noen evalueringsretoriske grep som har blitt gjennomført/gjennomføres på den utdanningspolitiske dagsordenen. Disse grepene kan sees på som strategier for å oppnå legitimitet til den offentlige styringen, og de kan muligens være med på å gi en

---

<sup>18</sup> Rapporten kom i 1988 og var et resultat av en evaluering gjennomført av OECD (Organisation for Economic Cooperation and Development).

<sup>19</sup> Prosjektet ble kalt EMIL (Det norske prosjektet om nasjonal evaluering og kvalitet i utdanningen). Se også 45.

forståelse for den utviklingen man i dag kan øyne på dagens utdanningspolitiske dagsorden.

#### 4.1. Teoretiske utviklingslinjer

Mange spørsmål kan stilles om hva som gjør at kvalitetsvurdering har fått en sentral plass på den utdanningspolitiske dagsordenen i dag. For å få svar på noen av spørsmålene vil det være nødvendig å se nærmere på den historiske og komparative konteksten. Forhåpentligvis kan man da få noen svar, se noen linjer og forstå dagens fokus bedre. Evaluering har ikke alltid vært en så betydelig del av skolehverdagen som den er i dag. Eller, man kan si at en type evaluering alltid har vært en del, men det er kun en snever type; testing.

I USA kan man se det første formelle forsøket på å evaluere skoleprestasjoner i Boston i 1845 (Madaus m.fl. 1993:5). Dette startet en lang tradisjon med elevenes testresultater som hovedkilde for å evaluere skolens effektivitet eller undervisningsprogram. Skriftlig eksamen ble nå introdusert, til fordel for muntlig som ble sett på som en urettferdig ordning fordi man der ikke kunne standardisere formen. En interessant bemerkning i tilknytning til denne evalueringen, er å vise til den skjulte politiske agendaen bak innføringen av skriftlige, standardiserte eksamener: Denne innsamlingen av data ble brukt til sammenlikning av skolene, som igjen ble brukt til å avgjøre rektors stilling. Her kan man se et tidlig eksempel på hvilken politisk styringsverdi evalueringsdata kan ha (på den skjulte politiske agendaen).

Madaus, Scriven og Stufflebeam har kalt perioden fra 1900 til 1930 for *The age of Efficiency and Testing* (op.cit:6). De beskriver denne perioden med stikkordene *systematisering*, *standardisering* og *effektivitet*. Metoder for å måle lærerne ble utviklet, tester som vurderte hvor effektive skolene og skolesystemene var ble brukt og man brukte forskjellige kriterier som forbruk, elev ”drop out”-rater og forfremmelsesrater for å vurdere effektiviteten. Testene ble sett på som objektive. Det problematiske med disse testmetodene var, i følge Madaus, Scriven og Stufflebeam, at man hentet inn eksterne evaluatorene som skulle avdekke feil og gi råd. I tillegg var det et problem at disse

”objektive” dataene ofte ble brukt som propaganda for å bygge opp demninger mot økende strømninger med offentlig kritikk (op.cit:7). Etter 1. verdenskrig økte bruken av standardiserte tester. Madaus, Scriven og Stufflebeam betegner dette som et stort problem, fordi det hersket en tro på at de ”objektive” dataene ga generaliseringsmuligheter (ibid).

Videre beskriver Madaus, Scriven og Stufflebeam perioden fra 1930 til 1945 som *The Tylerian Age* på grunn av ”the father of educational evaluation”, Ralph Tyler. Tyler hevdet at ”educational evaluation” skulle vurdere i hvilken grad de valgte målene ble nådd i et, eller som del i et, utdanningsprogram. Dette var tiden til den progressive utdanningsbevegelsen som hentet sin filosofi fra pragmatikken og adopterte verktøy fra behavioristisk psykologi (op.cit:8). Tyler var direkte involvert i denne bevegelsen og han ble headhuntet til prosjektet som man i dag kjenner som *The Eight Year Study*, 1932-1940. Prosjektet stilte spørsmål ved effektiviteten i den tradisjonelle high school. De ville kartlegge hvordan ferdighetene til denne elevgruppen var i forhold til ferdighetene til elevene i de progressive secondary skolene. Flere college hadde begynt å nekte progressive studenter adgang fordi de ikke nådde opp på poengskalaen i visse fag. *The Eight Year Study* introduserte et nytt og videre syn på evaluering i skolen enn det som hadde fokusert på effektivitet og testing. Tyler definerte evaluering som ”sammenlikning av intenderte mål med virkelige resultater” (op.cit:9). I lys av dette kan man si at evalueringen fikk et bredere fokus, man skulle vurdere resultater opp mot mål.

Perioden som kalles *The Age of Innocence* eller *The Age of Ignorance* (1946-1957) viser til en tid med lite interesse for evaluering (ibid). Lærere/utdanningsvitere trengte ikke å dokumentere effektiviteten i skolene. Evalueringsfeltet fikk en ny blomstringstid under *The Age of Expansion* (1958-1972). Etter USAs såkalte Sputniksjokk<sup>20</sup> i 1957 ble det satset for fullt på utdanning. I følge Madaus, Scriven og Stufflebeam startet dette en ny æra innen politisk, militær, teknologisk og vitenskapelig utvikling (op.cit:12). Med *The National Defense Education Act of 1958* utviklet det seg nye undervisningsprogram i matematikk, naturvitenskap og 2. språk, i tillegg til et utvidet fokus på veiledning og testing i skoledistriktene. Mange evaluatorene var skeptiske, og forespeilet at denne innsatsen ikke ville lykkes (ibid). Cronbach kritiserte evalueringen for å mangle

---

<sup>20</sup> Sputnik 1 – ”The world’s first artificial satellite” – var en russisk satellitt som ble sendt ut i rommet 4. oktober 1957. Dette startet ”romkappløpet” mellom USA og Sovjetunionen.

relevans og nytteverdi, og foreslo at evaluatorene måtte gå bort fra sin tilbøyelighet til å ”post hoc”-evaluere. Han mente at det var på tide å gå bort fra tilbøyeligheten til testing. Hans råd var å redefinere evalueringen og se den som en prosess av innsamling og tilbakemelding av informasjon som skulle hjelpe og guide curriculumutviklere (ibid).

I 1965 ble det lansert flere programmer som skulle bekjempe fattigdom. Programmene kan med en fellesbetegnelse kalles *The War on Poverty* (ibid). I stor grad handlet programmene om å skape like og bedre muligheter for alle amerikanske borgere gjennom blant annet utdanningssystemet. Målet var å skape det som Lyndon Johnson kalte ”The Great Society”. The Great Society legislation var med på å endre og utvikle utdannings- og samfunnsforskning (op.cit:13). Dette skulle vise seg å bli en betydningsfull periode for evalueringsvitenskapen. Gruppen initierte et stort utviklingsprosjekt: *The Elementary and Secondary Education Act of 1965* som ga mandat til å evaluere programmer for uheldig stilte studenter. Før 1965 betydde evaluering kun testing eller curriculum evaluering (House 1993:2). Etter 1965 strømmet forskjellige personer til for å tilby sine ideer om hvordan evaluering skulle gjennomføres. I tillegg til spesialister innen testing, kom det psykologer, sosiologer, økonomer, byråkrater og, til og med, noen filosofer. Selv om feltet var dominert av et psykometrisk forskningssyn (Madaus m.fl. 1993:14), kom nå nye og andre mennesker, ideer og problemer strømmende. Evalueringsteoretikere som Provus, Hammond, Eisner og Metfessel & Michael ville reformere Tylermodellen. Glaser, Tyler og Popham hevdet at kriteriebasert vurdering var et egnet alternativ til den normbaserte vurderingen (ibid). Cook ville bruke system-analytisk tilnærming. Han mente at man skulle evaluere mål, se på forutsetninger (input), undersøke implementeringen og kontrollere overleveringen, samt måle intenderte og uintenderte resultater. I tillegg understrekte han nødvendigheten av å bedømme hvor verdifull nytteverdien av evalueringsobjektet var (ibid).

I lys av denne utviklingen kan man hevde at evalueringsfeltet nå hadde utviklet seg til et tverrfaglig felt med separate tradisjoner og mange utviklingsretninger. Evaluatorene på utdanningsfeltet startet som oftest sin historiske tilnærming med Tyler, sin psykologiske tilnærming med Donald Campbell og sin sosiologiske tilnærming med Edward Suchman (House 1993:2). Det er rimelig å anta at en tidlig erfaring med testing og curriculum evaluering, ga utdanningsevaluatorene et forsprang i utviklingen, og man kan



muligens si at disse har vært blant de mest innflytelsesrike teoretikerne i det komplekse evalueringsfeltet.

Perioden som kalles *The Age of Professionalization* (1973 →<sup>21</sup>) startet med en identitetskrise i evalueringsverden (Madaus m.fl. 1993:15). Evaluatorene var ikke sikre på om de skulle være forskere, testere, administratorer, lærere eller filosofer, og i tillegg var de usikre på hvilke kvalifikasjoner de trengte. Madaus, Scriven og Stufflebeam beskriver evalueringsfeltet som fragmentert og usikkert, og det hadde, i følge dem, mistet sin legitimitet (ibid). Etableringen av mange evalueringstidsskrifter kan oppfattes som delaktig i på å profesjonalisere feltet. I tidsskriftene ble nye ideer og diskusjoner lagt merke til, og man kunne øyne en positiv utvikling. Universitetene tilbød nå vektall i evalueringsmetodologi, og noen utviklet til og med egne grader i evalueringsfaget. Med referanse til Madaus, Scriven og Stufflebeam kan man hevde at evalueringsfeltet nå beveget seg i retning av metaevaluering, noe som kan forstås som kvalitetssikring av feltet (op.cit:16).

I lys av dette er det mulig å si at evalueringsfeltet levde et intellektuelt og utviklende liv i mer enn et tiår. Kutt i samfunnsbudsjettet endret utviklingen (House 1993:2). Reagan-tiden ga få muligheter for evaluering, og feltet gikk inn i en svakere periode (ibid). I de stille 1980-årene kan man spore andre endringer også: Det kan se ut som om evaluering ble oppfattet som for viktig til at den skulle overlates kun til evaluatorene. Derfor utviklet store byråkratier sine egne interne evalueringsavdelinger (op.cit:3). Med den økende bruken av intern evaluering er det rimelig å anta at man ble utfordret av nye problemstillinger: Ovenfor hvem var de interne evaluatorene ansvarlig? Hvor offentlig skulle arbeidet deres være? Hvilke vitenskapelige kontroller var passende?

Det var ikke bare strukturen som forandret seg i denne tiden. Mellom 1965 og 1990 forandret metodologien, filosofien og politikken til evalueringen seg substansielt (ibid). House påpeker at evalueringsbegrepet nå beveget seg fra å være et monolittisk begrep til å bli et pluralistisk begrep, i tillegg til å favne multiple metoder, målinger, kriterier, perspektiver, grupper og interesser (ibid). Han viser hvordan evalueringen metodologisk

---

<sup>21</sup> Madaus, Scriven og Stufflebeam sier ikke noe om denne perioden er avsluttet. Man kan spørre seg om hvor lenge denne perioden varte. Med referanse til House kan man hevde at den endret seg etter ca. 10 år (House 1993).

forandret seg fra å kun være kvantitativ til også å favne kvalitative sider og hvordan den filosofisk forandret sitt syn på evalueringens verdi. Man kan beskrive dette ved å si at man gikk fra å mene at evalueringen var verdinøytral, til å se at det var flere verdihensyn man måtte forholde seg til. Politisk, hevder House, at evalueringens rolle utviklet seg fra å kun være en posisjon der evaluatorene så seg selv som tekniske og strategiske eksperter, til å innrømme at evaluering i seg selv hadde politisk interesse og effekt (op.cit:4).

#### 4.1.1. Et nasjonalt bilde

Som nevnt i innledningen til dette kapittelet, kan man iaktta en liknende utvikling i Norge. Haug har delt inn tidsperioden fra 1950 til 1990-tallet i tre reformperioder (Haug 1998). Han ser på perioden fra 1950 til 1970 som en tid med *Sentraliserte reformer*. Her hevder han at det finnes sterke bånd mellom departement, lærere og utdanningskomiteen, og at man kan se en kombinasjon av ovenifra og ned, pragmatisk og teknisk-instrumentell strategi. Evalueringens rolle blir av Haug karakterisert som sterkt kvantitativ, produktorientert og instrumentell. Haugs presentasjon av denne perioden kan muligens sammenliknes med perioden som overfor ble beskrevet som *The Age of Efficiency and Testing* (1900-1930), hvor standardisering og effektivitet var stikkord. Periodene kan ikke sies å være identiske, men man kan hevde at strategiene som ble brukt er sammenliknbare. Evaluering ble gjerne sett på som en måling som kun skulle testes opp mot målene. Dermed kan man si at strategiene var preget av en mål-resultat orientering. Det ble utviklet ”objektive” tester som kunne vurdere om resultatene var i samsvar med målene. I Norge er det kanskje ikke hensiktsmessig å bruke evalueringsbegrepet på perioden Haug har kalt *Sentraliserte reformer*. Det kan se ut som om den eneste formen for vurdering som var representert, var karaktersetting, eller resultatmåling av elevene. I sentrum for denne karaktersettingen stod Sandvæns skolemodenhets tester. Disse testene kan betraktes som psykometriske tester med en klar påvirkning fra amerikansk psykologi (Telhaug & Mediås 2003:151).

Den neste perioden Haug identifiserer strekker seg fra 1970 til 1990 (Haug 1998). Denne perioden, som han kaller *Indre reformer*, karakteriserer han med begrepene

desentralisering, idealisme og populisme. De store rammeplanene ga stor lokal frihet, og evalueringen spilte en minimal rolle, den var nesten fraværende (op.cit:10). Evalueringen i denne reformperioden betegner Haug som intuitiv, intern og kvalitativ, og ofte var det de samme personene som var ansvarlige for implementeringen som for evalueringen (ibid). Program som kan nevnes i denne sammenheng er EVA 1 (1974) og EVA 2 (1978)<sup>22</sup>, der EVA 2 lagde stor ståhei da forslaget om å fjerne karakterene på ungdomstrinnet ble fremmet. Etter dette var det rolige år på evalueringsfronten, tilnærmet ingenting skjedde på nasjonalt nivå (Telhaug & Mediås 2003:224-226). Det er vanskelig å finne en liknende periode i den amerikanske evalueringshistorien som skisseres av Madaus, Scriven og Stufflebeam. *Indre reformer* kan karakteriseres som en metodisk motsetning til perioden som ble kalt *Sentraliserte reformer* fordi det kvantitative fokuset nå ble erstattet med et kvalitativt fokus. Denne perioden representerte et alternativ til psykometriske prøver og fokuserte i stedet på prosess. Det kan derfor være rimelig å si at en formativ<sup>23</sup> evalueringstenkning var fremtredende.

Som nevnt ovenfor gav OECD i 1988 ut en tilstandsrapport på grunnlag av en evaluering av den norske skole. Hovedtema for evalueringen var hvordan de sentrale myndighetene kunne skaffe seg informasjonsgrunnlag for utformingen av utdanningspolitikken i en skole så sterkt desentralisert som den norske (Granheim 1990a). Denne evalueringen satte fart på skoleutviklingsprosjektene i utdanningssektoren. I rapporten sto det blant annet at Norge hadde svært lite dokumentasjon og styringsdata som kan si i hvilken grad målene nås. Rapporten spør: *"How can central authorities form an opinion of and influence the level of quality in a schoolsystem so strongly decentralized as the Norwegian one? The answer is to evaluate and monitor the system at different levels."* (OECD 1988). Rapporten slo fast at det er et sterkt behov for å utvikle en modell for evaluering av den norske skolen, der ulike nivåers rolle og ansvar blir klargjort.

Denne rapporten innledet en tredje, og siste, periode som Haug kaller *Rask, kompleks og politisk ad hoc-idealisme* (Haug 1998). Denne perioden kan beskrives med følgende stikkord:

---

<sup>22</sup> Evalueringsutvalget for skoleverket, 1972-1978.

<sup>23</sup> Formativ evaluering bør forklares i forhold til summativ evaluering. Scriven definerer summativ evaluering som "designing and using evaluation to judge merit", og formativ som "designing and using evaluation to improve the evaluand" (Shadish m.fl. 1991:73).

- idealistisk og intellektuell
- politisk makt og styring
- ovenifra og ned
- komplekse og omfattende endringer
- rask implementering (dempe all mulig motstand)
- debatter, endringer og modifikasjoner underveis i implementeringen, det vil si at beslutningsprosessen forflyttes og forlenges.

Man kan kanskje kritisere Haug for at denne perioden ikke fanger inn den utvikling som fant sted utover 1990-tallet<sup>24</sup>. Haugs siste periode ser ut til å kun ta for seg et ”moment” på den utdanningspolitiske dagsorden i denne perioden, nemlig Gudmund Hernes. Hernesreformene satte kanskje ikke evaluering i sentrum, men man kan hevde at det foregikk relativt store og betydelige evalueringsprosjekter både før og etter Hernes.

Kirke- og utdanningsminister Kirsti Kolle Grøndal startet et prosjekt på grunnlag av OECD-rapporten. Prosjektet het EMIL<sup>25</sup> og var svaret på en satsning der evaluering ble foreslått som et nøkkelement i arbeidet med å forbedre norsk utdanning. I 1990 avsluttet prosjektet med utgivelsen av boken *Utdanningskvalitet – styrbar eller ustyrlig*, redigert av Granheim, Lundgren og Tiller. Boken tar for seg to viktige problemstillinger: (1) De metodiske utfordringene vi står overfor når utdanningskvalitet skal måles, og (2) viktigheten av statistiske informasjonsbaser for utdanningssektoren (Granheim 1990a:10). Nå skulle departementet integrere arbeidet med å utvikle et nasjonalt vurderingssystem som del av reformarbeidet i grunnskolen og videregående opplæring. I de neste årene kan man se en forholdsvis stor aktivitet på evalueringsfronten, men bare på planleggingsstadiet, det vil si på nasjonalt nivå. Selv om man kan hevde at lite ble iverksatt eller implementert nedover i systemet, kan utdrag fra ulike politiske dokumenter vise at kvalitet og evaluering nå ble satt i søkelyset på den utdanningspolitiske dagsorden:

- *Evalueringssystemet må i størst mulig grad fange opp hele bredden i kunnskapsbegrepet selv om enkelte sider ved den samlede kompetansen er vanskeligere å måle enn andre.* St.Meld.nr.33 (91-92), s. 25.

<sup>24</sup> Merk at artikkelen det her refereres til er skrevet i 1998.

<sup>25</sup> Prosjektet ble også omtalt på side 38.

- *Utvikling av et nasjonalt vurderingssystem for grunnskolen er et virkemiddel for å heve kvaliteten på opplæringen for å ivareta enhetsskolens prinsipper.* St.prp.nr.1 (97-98).
- *Komme med forslag om til et helhetlig, nasjonalt vurderingssystem for grunnskolen ...* Moe-utvalgets mandat (Moe-utvalget 1997).
- *Kvalitetsutvalget.* Oppnevnt oktober 2001. Avga en delinnstilling i 2002 hvor det ble fremmet et forslag om å etablere et nasjonalt system for kvalitetsvurdering og oppfølging (NOU 2002:10, kap. 4).
- *Det skal være helhet, sammenheng og høy kvalitet i opplæringsløpet ... Utvalget skal beskrive, analysere og vurdere hovedtrekkene i dagens tilbudsstruktur, samt innhold, organisering og kvalitet i grunnopplæringen, i forhold til intensjonene i Reform 97 og Reform 94 ... Gjennom denne utredningen fremmer utvalget en lang rekke forslag for å bedre kvaliteten i norsk grunnopplæring* (NOU 2003:16, kap. 3.1 og 2.1).
- St.meld. nr. 30 (2003-2004) *Kultur for læring*. Meldingen er utarbeidet på bakgrunn av blant annet Kvalitetsutvalgets innstillinger, NOU 2002:10 og NOU 2003:16.

Denne fokuseringen kan muligens oppfattes som innledningen til en fjerde periode som vi kan se konturene av nå. Man kan hevde at evalueringsfokuset mellom 1992 og 1998 var preget av et reformfokus heller enn et fokus på kvalitetsvurdering. Man kan videre oppfatte dette reformfokuset som en type "ad hoc"-evaluering der man måtte starte på nytt hver gang. Ved å flytte fokus til kvalitetsvurdering kan man anta at ønsket er å skape et system som kan evaluere kontinuerlig. Dermed kan man si at man ønsker å bevege seg bort fra "ad hoc" og til institusjonalisering, det vil si at man ønsker å institusjonalisere et kvalitetsvurderingssystem. Hva et slikt fokus vil innebære, eller hva denne eventuelt nye perioden vil være preget av, vil jeg prøve å belyse videre i oppgaven.

## 4.2. Utvikling på den utdanningspolitiske dagsorden

Kapittel 2 prøvde å skape et bilde av at virkeligheten er konstruert og ikke alltid rasjonell. Når jeg har valgt å fokusere på symbol- og verktøyperspektivet for å belyse oppgavens problemstilling, kan man si at det er symboler som opererer i samfunnet rundt en organisasjon, og disses påvirkning på en organisasjon, som er i fokus. Likevel er det vanskelig å forstå disse påvirkningene uten å kartlegge den utviklingen som har foregått internt i organisasjonen. Det kan være problematisk å skille et ”utenfra og inn”-perspektiv fra et klassisk ”innenfra og ut”-perspektiv. ”Klassisk” brukes her fordi det som oftest har vært fokusert på dette ”innenfra og ut”-perspektivet når man har analysert organisasjoner (Røvik 1998). Det er derfor denne oppgaven vektlegger et ”utenfra og inn”-perspektiv da spørsmålet om *hvorfor kvalitetsvurdering har utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden* belyses. Ved å vektlegge dette perspektivet kan oppgaven muligens være med på å utvide en forståelse av *hvorfor-spørsmålet*. Oppgavens kapittel 5 fokuserer på dette perspektivet.

Selv om hovedfokus ligger på ”utenfra og inn”-perspektivet, er det som sagt nødvendig for oppgavens helhet å tegne et grovt bilde av den utviklingen, noen av de endringene, som har foregått internt, og som har betydning for kvalitetsvurderingsfokuset, på den utdanningspolitiske dagsorden. Et slikt bilde kan gjøre det enklere å forstå hvorfor ”utenfra og inn”-perspektivet blir sentralt, det vil si hvorfor det nå kan se ut som om den utdanningspolitiske dagsorden blir påvirket av flere og flere ideer og oppskrifter fra organisasjonssamfunnet.

Fremstillingen nedenfor er begrenset. Jeg har valgt å konsentrere fremstillingen om det sentrale og det regionale nivået da det er disse som er i fokus i oppgavens problemstilling. *Kommunene som aktører* kommer derfor ikke tydelig frem, og man kan dermed si at det er den politiske formuleringsarena som fremheves i dette avsnittet, ikke realiseringsarenaen (Lindensjö & Lundgren 1986). Dette valget er bevisst, det er verken tid eller rom for en slik utgreiing i denne oppgaven.

#### 4.2.1. Fra skoledirektørembeter til utdanningskontorer til ...

En måte å belyse disse endringene på – som kan sies å ha betydning for spørsmålet om hvorfor kvalitetsvurdering har blitt et sentralt element på den utdanningspolitiske dagsorden – er å beskrive utviklingen som har skjedd mellom det sentrale og regionale nivået. Dette kan betraktes som et spørsmål om hvem som skal styre den norske skole, altså et spørsmål om fordeling av makt og muligheter.

Endringene som beskrives her kan ikke sies å være ”særnorske”. Liknende endringer skjedde også utenfor Norges grenser, blant annet i Sverige og Danmark. Telhaug og Mediås viser i denne sammenheng til endringer som skjedde i Sverige under sosialdemokraten Göran Persson, ”en usedvanlig handlekraftig skoleminister” (Telhaug & Mediås 2003:395). De skisserer endringer i styringen av skolesystemet ved å vise til at Sverige omkring 1990 hadde kommet opp i en vanskelig økonomisk situasjon hvor behovet for å redusere de offentlige midlene var fremtredende. Ministeren uttalte på den sosialdemokratiske partikongressen høsten 1990 at han regnet med å fremme forslag for Riksdagen om å halvere den statlige skoleadministrasjonen. Han ville spare 100 millioner kroner på denne reduksjonen. Samtidig gjorde han et annet grep; han besluttet at hele personalet ved länskolenämndenes kontor skulle sies opp. Beslutningen ble formalisert i Riksdagen i desember 1990 og länskolenämndene ble avviklet medio juli neste år (ibid). Telhaug og Mediås ser på denne noe uvanlige handlemåten fra Perssons side som et forsøk på å ”unngå den mobilisering av særinteresser og ulike motkrefter som en tradisjonell handlingsmåte ville ha ført med seg” (ibid). Denne nedleggelsen av det statlig-regionale styringsleddet kan oppfattes som en konsekvens av en langsiktig og prinsipiell styringstenkning som fant det regionale leddet overflødig. Persson mente at den detalj- og regelstyrte skolen nå måtte byttes ut med en mål- og resultatorientert styring. I tillegg fant han det regionale leddet overflødig da han jobbet for en kommunalisering av lærerne som innebar overføring av makt til kommunene (ibid).

En liknende utvikling er mulig å finne i Danmark. Undervisningsminister og liberalist Bertel Haarder var ikke i tvil om at den danske skole måtte desentraliseres. Han argumenterte *mot* detaljstyring og *for* avtalesanering, en avbyråkratisering og disposisjonsfrihet på det kommunale og lokale nivået (op.cit:396). Hans forslag i 1989

var at mellomleddet mellom undervisningsministeriet og kommunene, amtskolekonsulentstillingen, skulle nedlegges. Stillingen forsvant i 1992.

Telhaug og Mediås hevder at de svenske og danske innvendingene mot den statlig-regionale styringen har mye til felles med argumentasjonen i Norge, selv om denne lenge artet seg på en litt annen måte. Uviljen mot skoledirektørembetet i Norge utspilte seg først og fremst på det kommunale nivået, representert av Norske Kommuners Sentralforbund og til en viss grad av kommunepolitikere (ibid). Argumentasjonen handlet om at *folket* selv skulle ta styringen og "bli herre i eget hus". Fra det kommunale nivået kan man si at skoledirektørembetet ble betraktet som en del av et statlig overformynderi som stod i veien for deres ønske om et kommunalt selvstyre. Det kommunale nivået hevdet at embetet fortsatt var en representant for den statlige styringen, selv om embetet hadde gjennomgått en utvikling fra legal og finansiell styring til ideologiske styring. Norske Kommuners Sentralforbund argumenterte for at kommunene ikke lenger trengte tilsyn, kontroll og veiledning. Kommunene hadde etter hvert utviklet en faglig administrasjon som selv kunne ta seg av slike oppgaver. I tillegg ble det fremsatt argumenter for at en avvikling av det regionale nivået ville spare staten for store utgifter.

Norske Kommuners Sentralforbund ble støttet av de største byene, men ut over det fikk ikke forbundet oppslutning om sine meninger. Den ble avvist av de politiske partiene, av Norsk Lærerlag og av den kommunale skolesektorens representanter (skolesjefene). Skolesektoren ønsket seg skoledirektøren som sakkyndig støttespiller i utviklingsarbeidet, og fra sentralt politisk hold ble embetet sett i et sosialt nasjonsbyggingsperspektiv hvor skoledirektørene skulle ivareta et nasjonalt ansvar og en nasjonal kunnskapspolitik i en tid som var økonomisk vanskelig også i Norge. De skulle fremme den nasjonale kunnskapspolitikken gjennom å sørge for at barn og unge fikk et likeverdig og felles opplæringsstilbud uavhengig av hvor i landet de bodde (ibid).

Man kan se på *St.meld. nr. 37 (1990-91) Om organisering og styring i utdanningssektoren* som en løsning på konflikten mellom kommunal og statlig styring. Meldingen, som kan sies inneholde en argumentasjon for målstyringsprinsippet, ga uttrykk for at det nå var behov for en klar og entydig statlig styring på områder av tydelig nasjonal relevans, og en samordning av utdanningspolitikken med tanke på god



sammenheng på tvers av utdanningsområdene. Regjeringen Harlem Brundtland, med Gudmund Hernes som utdanningsminister, viste først til den samordningstanken som sosialdemokratiet hadde holdt fast ved gjennom hele etterkrigstiden. Regjeringen hevdet at det nå var et behov for et organ på det regionale planet som kunne samordne utdanningspolitikken og koordinere offentlig utdanningsvirksomhet innenfor regionene på tvers av utdanningsområdene. Skoledirektørene hadde hittil kun hatt ansvar for grunnskolen, nå skulle det etableres et nytt organ som inkluderte videregående opplæring og høgre utdanning. Det regionale styringsleddet skulle også ha ansvar for å videreføre arbeidet med en rekke praktiske oppgaver som skoledirektøren hadde tatt hånd om som for eksempel lederopplæring, sentralt styrt forsøks- og utviklingsarbeid, etterutdanning og veiledning. Det regionale leddet skulle dermed i fremtiden delta i den ideologiske styringen av skoleverket. Regjeringen mente det regionale styringsleddet skulle være ansvarlig for en oppgave til, en oppgave som man kan hevde brøt med tidligere tenkning: I meldingen hevdet regjeringen at det var behov for en bedre og mer omfattende evaluering av virksomheten i skolen enn den skoledirektøren hadde utført. Det regionale leddet skulle nå registrere hvordan skolesystemet fungerte i alle ledd, og det ble også påbudt å ta ansvar for den praktiske gjennomføringen av resultatoppfølgingen og sørge for at resultatene ble formidlet til departementet (Telhaug & Mediås 2003:397-399).

Denne meldingen kan sees på som et retningsskifte i den statlige styringen. Kontrollen med vedtak som gjaldt rammefaktorene, ble svekket til fordel for en overvåkning som rettet seg mot de resultater som skolen oppnådde.

Meldingen fikk medhold i Stortinget. Som en følge av dette ble skoledirektørkontorene nedlagt og erstattet med statens utdanningskontorer på det regionale nivået. Vedtaket ble iverksatt 1. august 1992.

En av de klareste konsekvensene denne omorganiseringen fikk, var kanskje at de nye utdanningskontorene ble besatt med et lederskap (utdanningsdirektører) som ikke nødvendigvis hadde sin bakgrunn fra grunnskolen. Et sitat fra Telhaug og Mediås viser dette:

”Andelen av direktører med en allmennlærerutdanning innsiktet mot grunnskolen ble mindre. Tilsvarende økte innslaget av direktører som hadde utdanning for og praksis fra videregående opplæring og høyere utdanning. Det nye kollegiet fikk ikke bare flere medlemmer med hovedfagsutdanning i de tradisjonsrike norske lærdomsfagene, men det fikk også medlemmer med slik utdanning i fag som sosialøkonomi, administrasjon, elektronikk og utdanning for kjemiingeniører.” (op.cit:399).

Man kan hevde at et vesentlig trekk ved det nye kollegiet var at de samlet sett kjente grunnskolen dårligere enn det skoledirektørene gjorde, men at de hadde tilsvarende bedre kunnskap om høyskolesystemet. Denne forskyvningen i kompetanse kan sees på som en svekkelse av den pedagogisk teoretiske spisskompetansen, mens faglig eller skolefaglig kompetanse økte.

En annen endring kan også oppfattes som fremtredende: Det nye kollegiet ble preget av en større variasjon og bredde når det gjaldt praksis og erfaring fra samfunnslivet. De tidligere skoledirektørene hadde frem til 1992 vært tett knyttet til skolen som sektor, mange hadde erfaring, og mer erfaring enn de nye utdanningsdirektørene, fra arbeidet i de sakkyndige skolerådene og følgelig fra pedagogisk forsøks- og utviklingsarbeid. De hadde også deltatt sterkere i lærerorganisasjonene. Til gjengjeld kunne de nye utdanningsdirektørene vise til større deltakelse i politiske virksomhet, spesielt på det kommunale og fylkeskommunale planet (op.cit:400).

Intensjonene med omorganiseringen om å samordne grunnskole, videregående opplæring og høgre utdanning lot seg i liten grad realisere (op.cit:401). Statens utdanningskontorer hadde først og fremst med grunnskolen å gjøre. En endring som var mer fremtredende var at omorganiseringen førte til at det regionale leddet kom til å ivareta en mer likeartet og sterkere styring på vegne av de sentrale myndighetene. De nye utdanningsdirektørene ble dirigert med relativt hard hånd fra departementet noe som kan sees på et tegn på hvordan skoleutviklingen gjennom Hernes-Sandal-perioden ble drevet frem etter nasjonale målsettinger (op.cit:402).

Oppsummert kan man si at omorganiseringen i 1992 endret den statlig-regionale styringen særlig på tre områder. For det første endret kompetansen til ledelsen på det regionale planet seg, det var ikke lenger slik at direktørene stod for en kompetanse som var tett knyttet til grunnskolen. For det andre endret omorganiseringen forholdet mellom det statlige og det kommunale planet. Direktørene ble sterkere styrt fra sentralt hold,

departementets tildelingsbrev<sup>26</sup> ga en bredere og mer detaljert beskrivelse av direktørenes plikter enn det som hadde vært tilfelle før. Det tredje området som er verdt å merke seg kan sees på som en konsekvens av de to første. Direktørene i de nye utdanningskontorene var ikke, som skoledirektørkollegiet, kun knyttet til skolen som sektor, men viste seg som sterkere politiske aktører gjennom deltakelse i politisk virksomhet. Dermed kan man hevde at det utviklet seg et sterkere bånd mellom politikk og praksis, noe som viste seg i en sterkere kontroll og styring fra det sentrale leddet.

Smith-utvalget, som ble oppnevnt av regjeringen Harlem Brundtland i 1993, skulle gjennomgå lovverket for skoleverket med sikte på å gjøre det til et bedre og klarere styringsmiddel. Med denne oppnevningen kan man si at regjeringen bekreftet at flere tidligere arbeid hadde lagt opp til en overgang til større grad av mål- og resultatstyring (NOU 1995:18:24). Arbeidet endte i innstillingen NOU 1995:18 *Ny lovgivning om opplæringa* og kan betraktes som et arbeid som vektla desentralisering av rammer og videreføring av målstyringsideologien. I tillegg viste det seg nå et behov for en type *monitorering* av måloppnåelsen: ”Vektleggingen av mål- og resultatstyring gjennom rammeoverføringer til kommunesektoren skaper behov for tilpassede instrumenter for ”overvåkning” av måloppnåelse og resultatoppfølging. Særlig gjelder dette behovet for å sikre en tilnærmet lik og høy nasjonal standard på tilbudene.” (op.cit:54). Utvalget mente det var på tide å redusere detaljerte regler og standarder av virksomheten i lovgivningen og heller legge vekt på beskrivelser av hvilke funksjoner opplæringen skulle ivareta. Beskrivelsene skulle ligge til grunn for den nasjonale styringen, de skulle være nasjonalt utformede mål som skulle ”overvåkes”.

Smith-utvalgets arbeid kan sees på som et grunnlagsdokument for den nye opplæringsloven som kom i 1998. Den endelige lovbestemmelsen ble som følgende: ”Elevane har rett til nødvendig rådgivning om utdanning, yrkestilbod og yrkesval og om sosiale spørsmål. Departementet gir nærmare forskrifter.”<sup>27</sup>.

Med NOU 2000:22 *Om oppgavefordelingen mellom stat, region og kommune* ble den statlig-regionale styringen av grunnskolen utsatt for kritikk. Telhaug og Mediås henviser til oppgavefordelingsutvalgets arbeid:

<sup>26</sup> Departementets tildelingsbrev beskriver oppgaver og midler.

<sup>27</sup> *Lov om grunnskolen og den vidregående opplæringa av 1998 med endringar per 10. juni 1999, §9-2.*

Oppgavefordelingsutvalget, som regjering og Stortinget, kunne enes om at det kommunale selvstyret hadde blitt svekket som en følge av statens bruk av detaljstyrende virkemidler som regelverk, øremerkede tilskudd, handlingsplaner, plankrav, rapporteringskrav, statlig tilsyn og kontroll. Og dette hadde skjedd uten at det kunne dokumenteres at den statlige detaljstyringen hadde gitt et mer likeverdig tjenestetilbud. Den statlige detaljstyringen av kommunesektoren burde derfor reduseres og det kommunale handlingsrommet økes.” (Telhaug & Mediås 2003:403).

Oppgavefordelingsutvalget hevdet at en desentralisering av oppgaver og myndighet ville gi en mer effektiv bruk av ressursene, et mer likeverdig tjenestetilbud, en revitalisering av det kommunale demokratiet og økt brukerinnflytelse. Det er rimelig å anta at konsekvensen av en slik styringsfilosofi ville bli en endring i statens styring av grunnskolen. Kommunesektoren skulle nå få flere handlingsalternativ innenfor nasjonalt fastlagte rammer. Når det gjaldt det statlige-regionale styringsleddet i sin helhet, gikk både Oppgavefordelingsutvalget, regjeringen og Stortinget inn for at det skulle *effektiviseres* og *samordnes*. Det statlige-regionale nivået skulle særlig konsentrere seg om formidlingen av nasjonale mål, om veiledning angående lov- og regelverk, om behandling av klagesaker og kontroll av at lov- og regelverk ble fulgt (ibid).

Under Stortingets behandling av *St.meld. nr. 31* (2000-2001) *Kommune, fylke, stat – en bedre oppgavefordeling* fikk forslaget om at utdanningskontorene og fylkeslegen skulle integreres i fylkesmannembetet enstemmig tilslutning, men uten at utdanningskontorets oppgaver ble endelig avklart. Spørsmålet om utdanningskontorets oppgaver ble reist under budsjettbehandlingen i Stortinget høsten 2001. Her ble det opponert mot at de statlig-regionale oppgavene skulle begrenses til tilsyn, kontroll, klagebehandling, veiledning i forhold til lov- og regelverk og informasjon om nasjonale saker. Det ble argumentert for at utdanningskontorene fortsatt skulle ha en pådriverrolle (Telhaug & Mediås 2003:403). Flertallet i Kirke-, utdannings- og forskningsdepartementet mente at utdanningskontorene burde bruke mindre ressurser på kontroll og rapportering og i stedet prioritere utviklings- og veiledningsoppgaver. I *Budsjett-innst. S. nr. 12* (2001-2002:21) står det at det å gi kommuner og skoler hjelp med en målbevisst kvalitetsutvikling burde være en prioritert oppgave (op.cit:404). En slik filosofi kan muligens betraktes som en svekkelse av utdanningskontorenes funksjoner. Den kan kanskje også sees på som et ledd i en liberalistisk styringsfilosofi da det så ut til å gå

mot en avvikling av det statlig-regionale leddet som krevde resultater av de underordnede enhetene. Kan dette være med på å karakterisere en ny periode?

Telhaug og Mediås hevder at dette blir løsningen hvis Kvalitetsutvalget får det som de ønsker seg (Telhaug & Mediås 2003:404). Et hovedpremiss i innstillingen fra Kvalitetsutvalget er at skolens eier, den kommunale instans, må bevisstgjøres om sitt ansvar og settes i bedre stand til å lede grunnskolens utvikling. Det har også som premiss at det statlig-regionale leddet er beheftet med en rekke ulemper. Det legger beslag på betydelige ressurser og virker samtidig unødvendig byråkratiserende. Det avgjørende er kanskje likevel at utdanningskontorenes faglige miljø er for små og for svake. Derfor er et av forslagene til Kvalitetsutvalget at Statens utdanningskontorer avvikles (NOU 2002:10:34-36). Løsningsforslaget i innstillingen er at tilsyn og klagesaksbehandling overføres til fylkesmannen, mens høgskole- og universitetsmiljøene skal konkurrere om midler til utviklingsprogrammer som legges ut på anbud. Programmene skal etableres med egne brukerstyrever hvor representanter for eiere, arbeids- og næringsliv, foreldre, elever/lærlinger og lærere er med.

Forslaget kan altså føre til at det regionale leddet avvikles, noe som kan betraktes som en svekkelse av statens og den pedagogiske profesjonens styring av grunnskolen. *Kan dette sees på som en avideologisering av den utdanningspolitiske dagsorden?*

#### 4.3. Evalueringens retorikk – sentralisering eller desentralisering?

Kapittel 4.1 og 4.2 handlet om utvikling på to forskjellige felt, evalueringsteoretisk utvikling og utvikling på den utdanningspolitiske dagsorden. Her vil jeg prøve å se disse to feltene i forhold til hverandre og belyse evalueringens funksjon i et sentraliserings- og desentraliseringsperspektiv.

Kvalitetsvurdering kan sees på som et desentraliseringstiltak. "Skolen vet best" har vært et slagord på den utdanningspolitiske dagsorden<sup>28</sup>. Det lokale skal få større handlingsfrihet og skolene/skoleeier er selv ansvarlig for utviklingen og kvaliteten på

---

<sup>28</sup> Se blant annet [www.odin.dep.no/ufd](http://www.odin.dep.no/ufd) (rapporter, "Skolen vet best").

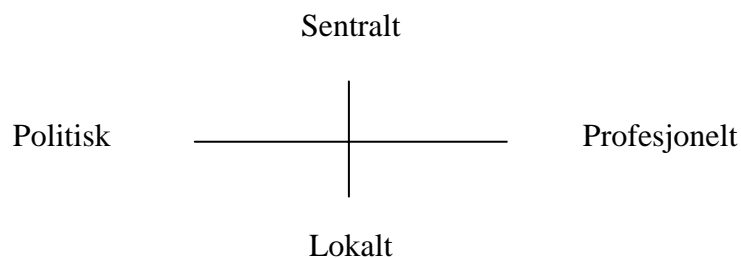
skolene. Weiler poengterer i tilknytning til dette hvordan desentralisering kan sees på som en ansvarsfraskrivelse fra statlig side. Han mener desentraliseringstiltak kan være et grep for å bygge opp legitimiteten til statlige utdanningsorganer. Følgende sitat kan belyse dette: "Decentralising the making and implementation of educational policy affects in very important ways the nature and interests of the state, particularly its ability to cope with the dual problems of policy conflict and erosion of its own legitimacy." (Weiler 1988/89:61) Weiler beskriver tre modeller for desentralisering som ligger bak forskjellige desentraliseringstiltak. Den første kaller kan "the redistribution model" (omfordelingsmodellen). Denne modellen indikerer en maktdeling. "The efficiency model"<sup>29</sup> (effektivitetsmodellen) handler om ressurser og utnyttelse av ressurser. Her er "cost-benefit"-analyser et verktøy. Modellen er tilpasset for å forsterke skolesystemets kostnads- og effektivitetsutregning gjennom en mer effektiv utnytting og forvaltning av tilgjengelige tjenester. Den tredje modellen er "the cultures of learning model" (lærekulturmodellen) som vektlegger en desentralisering av innholdet. Lundgren mener evaluering ofte kan sees på som en metode for å rekonstruere den sentrale rollen til en mildere, og mer liberal, styremåte, mens de tilegner de lokale myndighetene en strengere styring (Lundgren 1990). Et eksempel fra Oslo kommune kan belyse dette: I byrådets politiske grunnlag finnes det et eget punkt som omhandler ønske om å "... arbeide videre med å skille mellom bestiller- og utførerrollene og styrke bestillerrollen for å etablere klarere ansvarsforhold og skape mer effektiv organisering og sikre at etater og bydeler ivaretar befolkningens behov"<sup>30</sup>. Dette punktet, sammen med byrådets satsning på evaluering, kan man tolke som en strategi for å fordele ansvar utover i systemet, et retorisk grep for å skape et bilde av mildere sentral styring.

Desentralisering i utdanningssystemet setter søkelyset på balansen mellom politisk og profesjonell makt over utdanningen. Lundgren viser dette da han beskriver styring av utdanningssystemet i forhold til to dimensjoner (op.cit:33). På den ene siden er det et spørsmål om *hvor* makten er, og på den annen side er det et spørsmål om *hvem* som har makten. Han illustrerer dimensjonene i en modell (op.cit:34):

---

<sup>29</sup> Jamfør den blå bølge (NPM) – beskrives i kapittel 5.1.

<sup>30</sup> Se [www.oslo.kommune.no](http://www.oslo.kommune.no)



Forholdet mellom desentralisering og evaluering er etter min mening spennende da desentralisering kan oppfattes som en overføring av makt til det lokale, mens evaluering kan virke som en inndragelse av den delegerte makten. Hvem som styrer og hvem som har, og ikke minst hvem som tar, ansvaret, kan da forstås som et sentralt spørsmål i desentraliseringssammenheng. Lundgren tegner opp et bilde av forandringer han mener å se i utdanningssystemet. Han beskriver en bevegelse fra sentral styring til mer lokal styring og oppfatter dette som en skjerping omkring spørsmålet om forholdet mellom hvem som formelt *har* ansvaret og hvem som i realiteten *tar* ansvaret (op.cit:33). Dermed kan man si at en bevegelse mot desentralisering setter fokus på spørsmål rundt lærerprofesjonalitet og lærernes profesjonelle holdning til ansvar (ibid).

Et av hovedproblemene Lundgren skisserer er "... de utviskede grensene mellom sentrale og lokale nivå på den ene siden, og mellom politisk og profesjonelt ansvar på den andre... Et spørsmål som følger av dette kan da være: Hvordan kan et nasjonalt evalueringssystem fungere når det defineres forskjellige ansvarsområder?" (op.cit: 33-34).

Lundgrens tanker rundt desentraliseringsreformene i Norge på 80-tallet kan oppfattes som aktuelle i forhold til dagens fokusering på kvalitetsvurdering på den utdanningspolitiske dagsorden. OECD-rapporten fra 1988<sup>31</sup> og Lundgrens<sup>32</sup> tanker kan begge være med på å aktualisere for eksempel et dokumentasjonsfokus. Det er rimelig å anta at desentralisering i et offentlig skolesystem fordrer en eller annen form for kontroll, en kontroll som kan si noe om hvordan skolene klarer seg og om de arbeider etter læreplanen. OECD ga råd til det norske skolesystemet om å evaluere og veilede på

<sup>31</sup> "How can central authorities form an opinion of and influence the level of quality in a schoolsystem so strongly decentralized as the Norwegian one? The answer is to evaluate and monitor the system at different levels."

<sup>32</sup> "How can a national system operate by defining the various areas of responsibility?"

alle nivå i systemet (OECD 1988). Man kan dermed hevde at det er mulig å se en direkte link fra denne kartleggingen og teoretikere som Lundgrens poengteringer, til dagens fokusering på kvalitetsvurdering.

Evaluering på det nasjonale plan synes å gi retningslinjer som åpner for koordinering innenfor og mellom nivåer. De redskapene som blir brukt, kan oppfattes som en garanti for kontrollen av sentrale mål og verdier, og modeller for evalueringer innenfor systemet. Lundgren mener at ”problemet med dette ser ut til å være den offentlige debatten om skolen der evaluering blir tillagt ubegrensede muligheter, samtidig som oppfatningen av hva en konkret evaluering er, innskrenkes til enkelte prøver og kriterier for hva som er produktivitet og effektivitet.” (Lundgren 1990:40). Lundgren utdyper dette ved å vise til forskjellen mellom de to arenaene; politisk formulering og politisk realisering. Man kan si at de to arenaene opererer under forskjellige forhold og med forskjellige regler. Lundgren hevder at evaluering som et kontroll- og utviklingsredskap på formuleringsarenaen kan oppfattes som en løsning på ulike krav om et mer effektivt skolesystem, og kan dermed sees på som et verktøy for å kunne møte utfordringer som oppstår i et desentralisert system. På formuleringsarenaen kan evaluering tolkes som et generelt verktøy som kan være med på å skape en oppfatning om rasjonalitet, og dermed kanskje også oppnå legitimitet på den utdanningspolitiske dagsorden. Selv om mål blir formulert på en abstrakt måte, bør de sees på som produkter av konkrete og tunge forhandlinger, som etter hvert fremstår som solide, konkrete produkter (op.cit:40-41).

Lundgren påpeker videre at målene virker helt annerledes på realiseringsarenaen. Her vil det i følge han allerede eksistere konkrete mål. I denne sammenheng blir da mål og direktiver fra sentralt hold gjenstand for nye forhandlinger. På denne arenaen må evalueringer overføres til konkrete teknikker og metoder, understreker han. Og videre mener han at evalueringsverktøy som kommer ovenfra, dermed vil gå inn i en konkret læringsprosess og ha virkning på denne prosessens retning og struktur. Han fremhever dette ytterligere i følgende sitat: ”Omdanningen av evaluering til en konkret form blir da nøkkelspørsmålet til hvordan evaluering kan anvendes, og hvordan den vil svare til forventninger på arenaen for politisk formulering.” (op.cit:41).



Lundgrens utgreining om formulerings- og realiseringsarenaen kan gi assosiasjoner til både diskursteorien og symbolperspektivet. Med et diskursteoretisk utgangspunkt kan man hevde at det på formuleringsarenaen opererer ulike diskurser som kjemper en kamp for å oppnå hegemoni. De konkrete målene, og de valgte verktøyene, kan derfor sees på som diskurser som har vunnet kampen om å fiksere mening innenfor et bestemt felt. Man kan se dette som symboler som etter hvert sees som nødvendig for arenaen og derfor blir gitt rasjonalitet og ”oversatt” til konkrete verktøy som skal være med på å løse systemets utfordringer og oppgaver.

Kvale skriver i artikkelen *Kunnskapsevaluering og maktdesentralisering* at ”hvis en sentral kvalitetsvurdering/evaluering blir den avgjørende styringsfaktor i det fremtidige skolesystemet, kan desentralisering av budsjettkontroll og innlemming av kulturelle forhold godt bli en styring gjennom en illusjon av lokal deltagelse.” (Kvale 1990:113-114). Med dette mener han at man da kan oppleve en dobbel kommunikasjon, altså at det samtidig med en åpen overføring av juridisk og budsjettmessig kontroll fra det sentrale departementet til det lokale samfunnet bygges en sentral kunnskapskontroll gjennom en mindre synlig evalueringshånd inn i et nasjonalt evalueringssystem. Da kan man hevde at det er selve evalueringen som styrer utdanningen, og denne evalueringen på en måte skjuler sin styring, det vil si at den uttales som noe annet enn det den er. Evalueringen kan dermed ende opp som en skjult politisk funksjon (ibid).

Lundgrens refleksjoner omkring evaluering kan tolkes som en støtte til Kvale. Lundgren hevder at evaluering i seg selv kan bli en styrende faktor på den utdanningspolitiske dagsorden. Man kan kanskje si at det er vanlig å dele inn i tre sentrale styringsfaktorer: juridisk, økonomisk og faglig-profesjonell (Lundgren 1990). Lundgren har en oppfatning om at man kan legge til en fjerde styringsfaktor, nemlig evaluering (op.cit:39). Evaluering kan følgelig sees på som et verktøy som brukes på den utdanningspolitiske dagsorden for å styre utdanningssystemet i ønsket retning. Evalueringsresultatene kan da brukes som belegg for å gi ulike program rasjonalitet og for å oppnå legitimitet.

Det er rimelig å anta at det i fremtiden vil etableres nye styringsfaktorer. Man kan kanskje si at man i dag står overfor et stort, og noe uklart, styringsverktøy som omhandler *informasjon* og *kommunikasjon*. Det kan nok være kritikkverdig å kalle

informasjon og kommunikasjon et styringsverktøy. Det bør muligens heller sees på som en slags makt som er med på å produsere og påvirke den sosiale virkelighet på alle områder<sup>33</sup>. Man kan hevde at dette fører med seg en gjennomsiktighet (transparency) som det kan knytte seg ulike styringsmuligheter til.

Man kan se utviklingen mot et nasjonalt kvalitetsvurderingssystem i forhold til ideen om informasjon og kommunikasjon som styringsverktøy. Ved innsamling av utdanningsdata på nasjonalt nivå, kan det tenkes at man utvikler et verktøy som styrer gjennom åpenhet om prosesser og resultater (som måles på grunnlag av standardiserte kvalitetsindikatorer). Dermed er det mulig å anta at offentlighetens innsyn i skolehverdagen er med på å styre utviklingen på den enkelte skole i forhold til for eksempel ordninger med fritt skolevalg og diverse intensivordninger.

House hevder at en slik innsamling av utdanningsdata på nasjonalt nivå kan synes fornuftig, forutsatt at data blir håndtert på en forsvarlig måte (House 1990:152). Et problem med nasjonal innsamling av utdanningsdata kan være at de blir brukt i et forsøk på å kontrollere utdanningsresultatet heller enn som kilde til informasjon. Her kan man hevde at det er rimelig å trekke frem de flere amerikanske stater som eksempel. De har gjennomført et stort testprogram i forsøk på å påvirke utdanningsresultatene (ibid). House mener det er vanskelig å oppfatte programmene som kilde til å sikre informasjon, de bør heller tolkes som et forsøk på å etablere tallmessige indikatorer og standarder slik at lærere og skoler kan bli holdt ansvarlig. Han mener det å bruke obligatoriske ferdighetstester for å utøve påvirkning og tvang er en vanlig mentalitet blant lovgivere, styremedlemmer og sentrale administratorer. Denne tankegangen kan tolkes som frustrerte forsøk på å kontrollere utdanningsmessige forhold (ibid). House kritiserer disse programmene og påpeker at empiriske studier indikerer at tvangsoppleggene med testing til slutt slår feil. Han mener dette skjer fordi det viser seg at der man anvender strekt press, tilpasser lærerne seg med å undervise i teste (ibid). Hvis en slik utvikling blir en realitet, kan man betegne undervisningen med det som kan kalles "teaching to the test"<sup>34</sup>. House argumenterer for at det finnes gode grunner til at slik bruk av tester ikke fører til bedre utdanning i følgende sitat:

---

<sup>33</sup> Jamfør diskursteoriens maktbegrep, side 21.

<sup>34</sup> "Teaching to the test" kan forstås som at man underviser elevene i å bli gode til å gjennomføre standardiserte tester. Dette er et fenomen som er bedre kjent i USA enn i Norge. For videre lesing, se for

”Årsaksfaktorene bak bedre utdanning blir simpelthen ikke knyttet til indikatorer som testen måler. Det er som å prøve å forandre barometertrykket ved å manipulere barometeret. Den sanne årsaksfaktor ligger et annet sted. Likevel er det slik at avstanden mellom de faktorene som kan føre til forbedret utdanningskvalitet og til indikatorene i testen, fortsetter å unnsnippe sentrale myndigheters oppmerksomhet og bevissthet. Testresultater kan ikke med hell brukes som middel til å påtvinge utdanningsmessig forbedring. Nasjonale regjeringer har god grunn til å samle testdata for å ha et grunnlag for beslutninger, men det er noe helt annet.” (op.cit:152-153).

## 5. ARGUMENTER FOR KVALITETSVURDERING

Telhaug og Mediås identifiserer fire regimer med vekt på skolens kulturformidling, hegemoniske kunnskapsregimer og styringssubjekter i boken *Grunnskolen som nasjonsbygger. Fra statspietisme til nyliberalisme* (Telhaug & Mediås 2003). Det første regimet de skisserer strekker seg fra 1739 til 1850/60 og kjennetegnes av at det innholdsmessig speiler en europeisk kultur og styres av et *teologisk kunnskapsregime* og et *geistlig embetshierarki*. Dette regimet avløses langsomt av et *nasjonalhistorisk regime* hvor lærerne, sammen med lekfolk, er de mest innflytelsesrike styringssubjektene. Årene fra 1935 til 1980-årene danner et tredje regime. Her finner vi det klassiske og nyradikaliserte sosialdemokratiets folke-/grunnskole, hvor likhets- og likeverdstanken står i sentrum. Prosessene sidestilles med faglig substans som dannelsesmiddel og den *pedagogiske teorien* får sin storhetstid som kunnskapsregime. Frem mot årtusenskiftet ser vi konturene av et siste regime: *det siviløkonomiske regimet*. Nå betones skolens betydning for den nasjonale økonomien sterkere og mer ensidig enn tidligere, og pedagogikken taper sitt kunnskapshegemoni til fordel for siviløkonomisk kompetanse og siviløkonomiske perspektiver. Styringstenkningen i dette fjerde regimet preges av både-og<sup>35</sup>.

Dette siste regimet ”innebærer en påstand om at grunnskolen er på vei inn i et nytt regime som tidligere forskning har vært forsiktig med å definere.” (op.cit:443).

Det er dette ”regimet” jeg vil prøve å identifisere i de neste kapitlene av oppgaven. Det oppfattes problematisk for meg å snakke om et helt nytt regime da det er mange meninger og diskurser som kjemper i det diskursive felt. Telhaug og Mediås’ inndeling i regimer kan oppfattes som et forsøk på å tegne et ønsket bilde av den utviklingen skolen har vært gjennom. Deres inndeling, og deres betraktning av samfunnets tendenser, kan ikke sies å være *den korrekte*. Det vil alltid finnes forskjellige oppfatninger om hvordan virkeligheten er og har vært. Jeg kan likevel støtte forfatterens

---

<sup>35</sup> En både-og filosofi innebærer på den ene siden at de sentrale myndighetene formulerer klare, men ikke detaljpregede mål for grunnskolens virksomhet. De overordnede målene, gjerne angitt som brukernes rettigheter, beskrives så presist og entydig at det ikke hersker noen tvil om betydningen nedover i systemet. På den andre siden finner man en tydelig og reell desentralisering eller deregulering som gir de underliggende enhetene rikere anledning til å finne egne løsninger på de formulerte oppgavene (Telhaug & Mediås 2003:442).

forsøk på å ramme inn hvilke forandringer som har vist seg de siste årene innenfor skolesystemet. De belyser en utvikling som viser at det i dag er nye, og andre, influerende faktorer (symboler) som gjør seg gjeldende på den utdanningspolitiske dagsorden. Det er noe av det samme som er hensikten med dette kapitlet, nemlig å se på hvilke tendenser (diskurser) i samfunnet som influerer den utdanningspolitiske dagsordenen, og drøfte hvilke konsekvenser dette kan få for utdanningssystemets utvikling. Jeg vil hele tiden forholde meg til oppgavens problemstilling; *Hvorfor har kvalitetsvurdering utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden?*, og; *Hvilken argumentasjon legges til grunn for kvalitetsvurdering på den utdanningspolitiske dagsorden?* I tilknytning til problemstillingen synes det rimelig å diskutere hvordan et fenomen som kvalitetsvurdering eventuelt institusjonaliseres. Dette handler om hvilke diskurser som kjemper i samfunnet, og hvilke av disse som vinner frem på den utdanningspolitiske dagsorden. Man kan dermed se denne diskusjonen som en diskursiv kamp der enkelte diskurser, ved å bli gitt rasjonalitet og legitimitet, prøver å oppnå hegemoni, og eventuelt klarer dette og dermed institusjonaliseres i organisasjonen. Hvis dette lykkes kan man ut fra diskursteoriens termer hevde at diskursene objektiveres ved at man ”glemmer” at verden alltid konstitueres i makt og politikk<sup>36</sup>.

Kapittelet vil vise hvordan en ”ny filosofi” har entret det offentlige systemet og skapt et behov for forandring. Denne filosofien kan neppe hevdes å være en entydig oppskrift på god organisering, men den kan heller sees på som et samlekonsept som inneholder en rekke elementer som igjen er med på å endre det offentliges oppfatning av god og riktig organisering. Vetlesen skriver: ”... markedets verdier – økt produksjon, økt konsum, økt avkastning – er blitt opphøyd til tidens forbilde og dermed blitt bestemmende for samfunnslivet som sådan, ikke bare for den økonomiske sektor.” (Vetlesen 1998:202). Ut fra dette kan man muligens identifisere andre styringsfaktorer som spiller en avgjørende rolle på den utdanningspolitiske dagsordenen i dag, enn de man har sett tidligere. Rhyn belyser dette i følgende sitat:

”In sum, the situation is such that expectations placed on the education system and on schools are rising while available resources are on the decline and overall conditions are deteriorating. This is why greater efficiency and effectiveness are required wich presupposes administrative reforms. ... The key word is NPM.” (Rhyn 2003:8).

<sup>36</sup> Diskursteoriens forståelse av begrepene makt og politikk, se side 21.

## 5.1. New Public Management-filosofien

New Public Management<sup>37</sup>-filosofien, den responsive stat og den refleksive modernitet er tre samfunnsfenomener som kan sies å spille en rolle i dagens politikkutforming, og dermed være med på å utforme den utdanningspolitiske dagsordenen. De kan alle tre være med på å gi en videre forståelse av *hvorfor* kvalitetsvurdering har utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden, og for hvordan man argumenterer *for* kvalitetsvurdering på den utdanningspolitiske dagsorden.

NPM kan betraktes som en samlebetegnelse for den filosofien som har preget den offentlige sektoren i den vestlige verden siden 1980-tallet. Klausen hevder at utgangspunktet for NPM er en kritikk av det offentlige for å være for stor, for byråkratisk i negativ betydning, det vil si regelfiksert og ufleksibel, og med for dårlig ledelse og for svak brukerorientering. NPMs løsningsforslag er å lære av det private (Klausen 2001).

I en artikkel skrevet til Nytt Norsk Tidsskrift i 1986 karakteriserer Olsen tre alternative fremtidsvisjoner som kan tolkes som utfordringer til forhandlingsstaten<sup>38</sup> og dermed til den nordiske velferdsmodell (Olsen 1986). De tre alternativene blir av Olsen karakterisert som tre bølger: en rød – som hviler på sosialistiske/solidariske verdier og et oppgjør med den herskende orden, en grønn – som skjerper oppmerksomheten om økonomisk og økologisk bærekraftighet, og en blå – som representerer en nyliberalistisk og nykonservativ orientering som krever mer individualisme, konkurranse og bruk av markeder (op.cit:9). I lys av disse bølgene tolker Klausen den blå bølge som representativ for NPM, som han oppfatter som et samlekonsept for flere delelementer og teknikker som har et felles utgangspunkt: inspirasjonen stammer fra det private og/eller fra en økonomisk og rasjonell tankegang (Klausen 2001:25). Videre hevder Olsen at de tre bølgene kan plasseres i hver sin periode. Den røde bølgen stod sentralt fra slutten av 1960-tallet til slutten av 1970-tallet. Denne bølgen overlappet den grønne

---

<sup>37</sup> Heretter omtales New Public Management med forkortelsen NPM.

<sup>38</sup> Olsen skriver at hovedtendensen i etterkrigstiden har vært styrking av forhandlingsmodellen, på bekostning av de andre styringsmodellene (den samfunnsstyrende stat, staten som supermarked og staten som moralsk fellesskap). I denne styringsmodellen problematiseres forestillingen om statens maktmonopol og antagelsen om at staten er sentralisert, samordnet og klart differensiert fra andre samfunnsorganisasjoner (Olsen 1986:6-8).

bølgen som ble satt på dagsorden på begynnelsen av 1970-tallet og mistet sin styrke som reformprogram i 80-årene. Den blå bølgen ble introdusert på midten av 1980-tallet og ble styrket på 90-tallet da det ble iverksatt en rekke reformer som hadde til hensikt å effektivisere den offentlige sektor i en liberalistisk reform (Olsen 1986:9).

Klausen ser som sagt likheten mellom Olsens blå bølge og NPM. Han forklarer NPM med utgangspunkt i to søyler som begge stammer fra det private (Klausen 2001). Den første søylen er den økonomiske argumentasjon og innebærer et ønske om en liberalistisk markedsorientering i og av det offentlige begrunnet i økonomiske resonnementer. Dette ønsket kommer til uttrykk i forestillinger om det hensiktsmessige i bruken av en rekke instrumenter som: privatisering, konkurranseutsetting og kontraktstyring, frie forbruksvalg, brukerbetaling, profittsenter-tanker, inntektsdekket virksomhet, selektiv lønnsincitament og incentiver. Den andre søylen omhandler det organisatoriske og ledelsesmessige. Her tenker man at det er fordelaktig å overføre kjente og utprøvde prinsipper fra det private til det offentlige, og dermed sette fokus på: entreprenøraktig ledelse, strategisk ledelse, teamledelse, samt en rekke konsepter som Service Management, TQM<sup>39</sup> og Business Process Reengineering, og til slutt mål- og rammestyring, nye økonomistyrings-, regnskaps- og budsjettssystemer.

Man kan muligens her hevde at den første søylen bygger på en markedsmessig styringslogikk, der aktørene forventes å tilpasse seg ut fra en vurdering av konsekvensene for egen nytte, og der man antar at konkurransemekanismen vil være attraktiv. Dermed kan man tolke den blå bølge som en representant for en styringsform som beveger seg mot *resultatstyring*.

Busch analyser av NPMs teoretiske grunnlag kan være med på å belyse NPM videre. Han mener dette grunnlaget må studeres ut fra hvilken rasjonalitet som ligger til grunn (Busch 2001). Han skriver i den sammenheng om tre viktige teorier som han oppfatter som grunnleggende for NPM. Den første er prinsipal-agent-teorien hvor det fundamentale er at en *prinsipal* er avhengig av handlingene til *agenten* for å oppnå egne mål. Den andre teorien han beskriver som viktig for NPM, er Public Choice-teori som er et forskningsfelt som studerer økonomisk atferd innen offentlig sektor. Den tredje

---

<sup>39</sup> Total Quality Management

teorien er transaksjonskostnadsteori hvor valg mellom alternative kontraktsformer og transaksjonskostnader er sentrale elementer.

Buschs tre grunnleggende teorier peker alle på at opportunisme<sup>40</sup> kan være et problem i offentlig sektor (op.cit:88). NPM formulerer større vektlegging av *konkurranse* som en av løsningene til dette problemet (ibid). Dette innebærer at det offentlige:

1. vektlegger resultatkontroll,
2. vurderer om resultatene er på et akseptabelt nivå i forhold til andre produkter/tjenester og
3. gir kjøperen mulighet til å vurdere om det som leveres, er verdt prisen.

Dette er viktige forutsetninger som bør være til stede for at konkurranse skal være et velegnet virkemiddel, hevder Busch. Han sier videre: "Det må være mulig å måle resultater, det bør være et marked som fungerer (flere aktører), og kjøperen må være i stand til å vurdere hva som leveres." (ibid). Man kan muligens hevde at konkurranseaspektet vil miste sin effekt dersom det svikter på et av disse punktene.

Hvis det ikke er mulig å måle seg med den private sektoren for å skape konkurranse, foreslår NPM at det kan settes i gang andre konkurransefremmende tiltak (ibid). Man kan fremme *intern konkurranse* ved for eksempel å innføre benchmarking<sup>41</sup>. Her er formålet er å sammenlikne offentlige virksomheter for å avdekke hvem som har de beste resultatene, og å avdekke styrker og svakheter i en virksomhet.

Det er også mulig å etablere *incentivsystemer* som et konkurransefremmende tiltak. Incentivsystemer kan skape en konkurransesituasjon mellom ansatte ved at man blir straffet eller belønnet for sin innsats.

Et tredje alternativ er å skape en *konkurranse om brukerne*. Dette kan man gjøre gjennom strategier som fritt konsumvalg, som for eksempel kan være fritt skolevalg, og brukerbetaling.

---

<sup>40</sup> Opportunisme: Individ handler i egen interesse, men handling i egeninteresse kan gå mot interessene til egen organisasjon og avtaleparter.

<sup>41</sup> Benchmarking innebærer at man systematisk sammenlikner seg med konkurrenter (og andre) på en rekke nøkkelindikatorer, og derigjennom får stilt i klarsyn forskjeller og forbedringsmuligheter (Røvik 1998:257). Mer om benchmarking som strategi i kapittel 5.5.2.



Det er også en mulighet å *konkurrere om ressurser*. Et eksempel på dette er en stykkprisfinansiering der en skole mottar et fast beløp per elev per år. Et annet eksempel kan være benchmarking knyttet til ressursfordeling hvor virksomheter får tilført midler ut fra sin produksjon og den angitte prisen.

## 5.2. Evalueringsbølgen

I et institusjonelt perspektiv med røtter i organisasjonsteori ser man evaluering som et organisatorisk formelement, det vil si i samtiden populære, sosialt legitimerede og utbredte ideer som gir oppskrifter på hvordan avgrensede deler av virksomheten bør organiseres, en "oppskrift" på lik linje som andre produkter som for eksempel lønnsystemer (Røvik 1998:279).

Dahler-Larsen hevder at det i dagens organisasjonssamfunn er legitimt å snakke om en evalueringsbølge som påvirker måten alle snakker om, og driver, organisasjoner på (Dahler-Larsen 2001a:11). Han illustrerer denne bølgen med flere eksempler i *Den Rituelle Refleksjon*:

- Evalueringer har gått fra å være engangsforeteelser til å bli rutinemessig anvendelser.
- Evalueringsfenomenet har ofte de egenskaper som kjennetegner en institusjon: det har blitt en del av vår kultur og er med på å strukturere hva vi snakker om, hva vi synes er vesentlige problemer og hva vi har tenkt å gjøre med de problemene.
- Evalueringens seks mysterier: (1) Hvorfor er nettopp evaluering blitt utbredt i vår aktuelle samfunnstype? (2) Den er et kollektivt oppfattet imperativ; alle vet at man skal evaluere, men ikke hvorfor, og ikke hva den skal brukes til, (3) Evalueringer settes i gang uten klare kriterier, sier ingenting om kvaliteten på ytelsene (mangler både programteori og kriterier), (4) Lytter til omgivelsene (særlig kundene) gjennom evalueringer (den responsive stat), (5) Vi kan se mye evaluering, men lite bruk av evaluering, og (6) Hvis ikke den instrumentelle

bruken er drivkraften, hva er det da? Hvilke drivkrefter utover de instrumentelle ligger bak evalueringsbølgen?

- Evaluering må i dag sees på som ”protected discourse”<sup>42</sup> som best lar seg forklare gjennom en normativ orientering av evalueringer som prosedyrer og ritualer. De er symbolske handlinger og har sin egen logikk. Med dette mener Dahler-Larsen at evalueringer fører til noe av seg selv, inklusivt effekter som ingen har tenkt på forut for evalueringen.
- Evalueringer er et grensefenomen, de er ”liminale tilstander”<sup>43</sup>. Evaluering er en situasjon der man stanser opp og reflekterende betrakter sine erfaringer midt i en bestemt sosial praksis. Dahler-Larsen henviser til Schultz som mener at evalueringer likner det som mennesket i videste forstand foretar seg når det skaper mening. Evalueringer er fortolkningsmessig ”modne” fordi de identifiserer en bestemt sosial praksis (det evaluerende program) og de valoriserer denne praksisen (tillegger den verdi). Evalueringer er anledninger til nyvurderinger og fortolkninger. De er ”modne” med mening. De er samtidig ”liminale” da de både vedrører systematisk vitensdannelse på den ene siden og politisk handling på den andre.

Denne evalueringsbølgen kan beskrives som et ritual som er med på å bekrefte og styrke bestemte tolkninger av virkeligheten, og stille seg i veien for andre. Dette ritualet kan dermed være med på å bestemme hvordan man skal forholde seg til, og forstå, virkeligheten, det vil si at det er ritualet som bestemmer hvilke diskurser som skal fremheves i det diskursive felt.

Foss Hansens beskrivelse av evaluering som et stort formelement<sup>44</sup> kan forstås som en ytterligere markering av evalueringsbølgen (Foss Hansen 2001:110). Hun reflekterer i artikkelen *Rejsende praksis: Hvilke forhold påvirker utviklingen af evalueringspraksis*

---

<sup>42</sup> Dahler-Larsen henter dette begrepet fra James March, og det kan beskrives som noe som holdes hemmelig, noe man ikke stiller spørsmål ved.

<sup>43</sup> Liminale tilstander er et begrep fra antropologien, og det betegner spesielle opplevelsesformer, for eksempel når man kommer ut fra og er på vei inn i en ny kultur. Det kan også være ritualer innenfor en kultur, hvor man kommer i kontakt med noe annet enn det dagligdagse. Det er overgangssituasjoner, eller spesielle tilstander, hvor man både er innenfor og utenfor. Det er anledninger til kritiske spørsmål og ettertanker. Det er tilstander som både kan true og styrke samfunnets strukturer, og som kan gi innsikt i samfunnets sentrale verdier (Dahler-Larsen 2001a).

<sup>44</sup> Formelement kan beskrives med Røviks definisjon av ”institusjonaliserte formelementer”. Han definerer disse som ”(i samtiden) populære, sosialt legitimerede og utbredte ideer som gir oppskrifter på hvordan avgrensede deler av virksomheten bør organiseres” (Røvik 1998:279).

på forskjellige politikområder? over de mange forskjellige måtene å designe og gjennomføre evalueringer på, og viser til at litteraturen om evaluering tilbyr en vifte av evalueringsmodeller (ibid). Disse kan betraktes som det Foss Hansen kaller en verktøykasse som bestillere, evaluatore og andre interessenter kan designe konkrete evalueringer på grunnlag av. Denne litteraturen kan derfor også anvendes som referanseramme for analyser av karakteren og utviklingen av evalueringspraksis på forskjellige politiske områder. Den kan videre hjelpe til å synliggjøre variasjoner og endringer i evalueringspraksis, samt til å synliggjøre mulige, men ikke benyttede, alternative anvendelsesformer (ibid).

Man kan finne mange slike ”modellvifter” i litteraturen. Et av mange eksempler er Vedungs modeller som beskriver forskjellige organiserende prinsipper for evalueringer<sup>45</sup> (Vedung 1998).

### 5.2.1. NPM, evaluering og den responsive stat

Fokuseringen på evaluering kan sees som en del av NPM. I tillegg kan man kanskje hevde at NPM kjennetegnes ved *den responsive stat*, som bygger på en forståelse av at staten skal innrettes etter de enkelte individers behov og preferanser. Videre kan man si at NPM kjennetegnes ved at fokus flyttes fra prosess til produkt, samt at utgiftene knyttes til outputs som kan måles med kvantitative indikatorer, noe som man igjen kan hevde at fremmer konkurranse.

---

<sup>45</sup> Vedungs modeller beskriver forskjellige organiserende prinsipper for evaluering. Forenklet kan modellene beskrives slik (Vedung 1998:49-90):

- Måloppfyllelsesmodellen: Det organiserende prinsipp er den evaluerende innsats' egne mål. Er disse oppfylt, og skyldes måloppfyllelsen i så fall innsatsen?
- Effektmodellen: Fokuset her er hvilke effekter innsatsen har resultert i, intenderte så vel som uintenderte.
- Prosessmodellen: Evalueringen fokuserer på implementeringsprosessen.
- Brukermodellen: Fokuset er på brukernes vurdering av innsatsen.
- Interessentmodellen: Her fokuseres det på alle interessentenes vurderinger av innsatsen.
- Økonomisk modell: Dette er en samlebetegnelse for evalueringer foretatt med fokus på forholdet mellom input og output/outcome, for eksempel cost-efficiency, cost-effectiveness eller cost-benefit undersøkelser.
- Kollegial modell: Modellen innebærer at det er fagfolk som vurderer fagfolk, for eksempel leger som vurderer andre legers innsats.

Dahler-Larsen setter søkelyset på sammenhengen mellom vår tids evaluerings-boom og NPM. Han poengterer hvordan denne filosofien i økende grad inspirerer den offentlige sektor, og ser at resultatorientering peker seg ut som et hovedtema i NPM (Dahler-Larsen 2001a:10). Her oppstår det imidlertid et problem, hevder Dahler-Larsen. Han mener det blir problematisk for den offentlige sektor å følge NPM-filosofien da denne sektoren ikke styrer etter de samme vilkårene som markedet gjør. Den offentlige sektor må dermed utvikle kunstige uttrykk (symboler) for kvalitet og kvantitet for sine ytelser. Lane støtter denne forståelsen av filosofiens inntog i den offentlige sektor (Lane 1995). Han mener man innenfor NPM *tror* at evaluering og effektivitetsmålinger er de riktige måtene å styre offentlig virksomhet på (op.cit:170). Dahler-Larsen mener i denne sammenheng at NPMs ideer gjør den offentlige sektor mer markedsorientert, i betydningen responsiv overfor de ønsker og krav som (for-)brukergrupper stiller. Han viser dette ved å se på det han kaller ”administrasjonspolitiske nøkkelord”, og hevder at man kan se en linje fra først strukturutvikling og senere resultatorientering, målstyring og ”firtatisering”. I den siste tiden mener han det er ordet ”dialogstyring” som har fått en sentral plass (Dahler-Larsen 2001a:10). Tanken bak denne utviklingen kan oppfattes som et ønske om å styrke de offentlige systemenes evne til å levere det de skal ved å intensivere kommunikasjonen mellom serviceproduserende enheter og dels politikere, dels brukere. Et ønske som dette kan tolkes dit hen at offentlig virksomhet opptrer som den responsive stat. Hva dette består i, kan muligens avklares i en dialog som formidler klare politiske målformuleringer, varedeklarasjoner på offentlig service, og ikke minst, informasjon om brukernes tilfredshet med det som faktisk leveres. Det å lytte til brukerne ved hjelp av undersøkelser virker dermed ikke bare som et redskap for å evaluere utvalgte programmer, men kan sees på som et integrert element i selve den administrasjonspolitiske tenkningen (ibid).

Dermed kan man hevde, med referanse til Lane, at utbredelsen av evalueringer slår igjennom i organisatoriske og politiske tenkemåter og regimer (Lane 1995). Dahler-Larsen utdyper dette ved å vise til at når man i dag registrerer et fokus på tidstypiske nøkkelord som kvalitetssikring, målstyring og resultatstyring, ser man at evaluering, eller noe som likner, spiller en avgjørende rolle. Evaluering, som tidligere var engangsforeteelser som skulle kartlegge et utvalgt programs grad av måloppfyllelse, har i dag utviklet seg til å bli brukt mer rutinemessig. Evalueringer er dermed i ferd med å integreres i administrative rutiner på mange nivåer og i forskjellige institusjoner

(Dahler-Larsen 2001a:10-11). Vedung mener konsekvensen av det som Dahler-Larsen betegner som en evaluerings-boom, er at kvalifikasjoner innenfor evalueringsfeltet i dag blir obligatorisk for profesjonelle i den offentlige sektor (Vedung 1998).

I følge de nevnte antagelsene ovenfor kan man si at evaluering som sådan har fått en betydelig plass i organisasjonssamfunnet; evalueringer kan i dag sees på som en viktig brikke både i organisering og styring av organisasjonene. Ovenfor ble det henvist til Dahler-Larsen og hans oppfatning av at evalueringsfenomenet har alle de egenskaper som kjennetegner en institusjon (Dahler-Larsen 2001a). Han mener at evalueringer har blitt en del av vår kultur og er med på å strukturere hva man snakker om, hva man mener er vesentlige problemer og hva man planlegger å gjøre med disse problemene. Videre karakteriserer han dette som en mani eller en slags kollektiv, mental besettelse (op.cit:11).

Ettersom evaluering her forstås som en mani, kan man kanskje oppfatte den som en enkel og ukomplisert løsning på dagens utfordringer i organisasjoner. I følge Dahler-Larsen og Krogstrup er dette ikke tilfellet. De hevder at det er et komplisert arbeid å avdekke hvilke elementer som influerer i en evaluering og dermed konkludere på grunnlag av evalueringsdata alene (Dahler-Larsen & Krogstrup 2000). En evaluering kan påvirkes av ulike elementer som kan sies å ha betydning for utfallet av evalueringen. Derfor bør en evaluator være oppmerksom på en rekke momenter som kan få betydning for utfallet (Dahler-Larsen 2001a). Blant annet hevder Dahler-Larsen at en evaluator må forholde seg åpent og aktivt til organisasjonen og ikke fokusere på et isolert program alene. Dette kan tolkes dit hen at evaluator må se sitt program i forhold til en institusjonell sammenheng. Evaluator må også, i følge Dahler-Larsen, forholde seg kritisk til ulike styringsmekanismer som kan påvirke evalueringsprosessen. Man kan kanskje se på det som en oppfordring til at evaluator bør opptre som en kritisk aktør under hele evalueringsprosessen, og ikke spille på lag med deler av organisasjonen. En siste oppfordring fra Dahler-Larsen er at evaluator alltid må huske på at all form for evaluering impliserer valg av verdier (op.cit:164-167).

### 5.3. Kvalitetsbølgen

Bruken av begrepet kvalitet kan oppfattes som en ”bølge” i tiden. Man kan hevde at begrepet brukes flittig på den utdanningspolitiske dagsorden og at det har utviklet seg til et mantra i utdanningspolitikken. Årsaken til denne fokuseringen kan være at begrepet oppfattes som et retorisk fordelaktig begrep.

I avisartikkelen *Kvalitet, kvalitet, kvalitet* intervjuer journalist Ola Henmo filosof Knut Ågotnes ved Universitetet i Bergen (Henmo 2003). Ågotnes reflekterer over bruken av begrepet kvalitet, og sier: ”Ved å påberope seg å være noe alle må være enig i, trekker betegnelsen oppmerksomheten bort fra innholdet i det tiltaket som skal settes i verk.” Dette kan tolkes dit hen at den utdanningspolitiske dagsorden kan snakke om kvalitet uten å ha noen virkelig strategi for hvordan denne kvaliteten skal utvikles. Henmo mener at den utdanningspolitiske dagsorden har fattet et godt retorisk valg ved å satse på kvalitet. Han oppfatter begrepet som et fenomenalt fleksibelt begrep som nærmest ikke har noen retoriske ulemper (ibid).

Lundgren blir også intervjuet i denne artikkelen. Han mener at desto oftere politikerne sier ”kvalitet” desto mer usikre er de på hva de snakker om. Lundgren understreker at begrepet får en rekke retoriske fordeler da ingen kan si de er imot kvalitet, og han sier: ”Og trenden er internasjonal. ‘Kvalitet’ brukes som et slags mantra i skoledebatten over hele kloden. Det er et sikkert tegn på at politikerne egentlig famler i blinde.” Lundgren hevder at kvalitetsfokusering kommer som et forsøk på kontrollere og skape et bilde av at den utdanningspolitiske dagsorden har klart for seg hvordan fremtidens gode skole skal se ut. Men han poengterer at fremtiden er usikker og at ingen kan si hva som kan være, og er, den gode skole i fremtiden. I dag er det vanskeligere enn noensinne å si hvilken kunnskap som er nødvendig i fremtidens samfunn (ibid). Lundgren mener politikerne prøver å vise handlekraft gjennom å måle og kontrollere mest mulig. Han påpeker at de kun måler forhold vi mener er viktige i dag. Om disse testene måler noe relevant for fremtiden, er det ingen som vet (ibid).

Tiller understreker faren ved å bli for opphengt i målinger i artikkelen ”Kvalitetsskring i hverdagen – Hva med grunnleggende verdier?” (Tiller 1990). Han mener vi kan bli så

opphengt i målinger og tester at vi glemmer å vurdere kvaliteten på de feltene som ikke kan måles, men som må være gjenstand for det gode skjønn.

### 5.3.1. (E)quality

Kvalitetsbegrepets inntog på den utdanningspolitiske dagsordenen bør, etter min mening, sees i sammenheng med utvikling og endringer i samfunnet<sup>46</sup>. Man kan hevde at det har skjedd en begrepsmessig utvikling på den utdanningspolitiske dagsordenen. I etterkrigstiden kan man si at likhetsbegrepet (*equality*) har vært et av de mest sentrale begrepene på den utdanningspolitiske dagsordenen. Likhet har vært et politisk mål i utdanningen, både når det gjelder formallikhet, mekanisk likhet, resultatlikhet, prinsipiell likhet og resultatlikhet. Nå har begrepet kvalitet (*quality*) kommet i fokus på dagsordenen (Lundgren 2003). Man kan kanskje ikke med rett si at *equality* er borte i utdanningspolitikken selv om *quality* har utviklet seg til å bli et sentralt begrep, men man kan se en endring i fokus fra likhetstanken til fokus på kvalitet og effektivitet. Dermed kan man muligens hevde at det i dag er kvalitet som er *nodalpunktet*, altså det tegnet som har privilegert status i forhold til andre tegn på den utdanningspolitiske dagsordenen. Man kan i dag betrakte bruken kvalitetsbegrepet nærmest som et økonomisk begrep som fremmer konkurranse og effektivitet, og ikke minst evaluering. Lundgren betegner dette i følgende sitat: "During the 1980s, and still more in the 1990s, the *quality* of education came in focus. It was as the prefix *e* in *equality* was dropped. The *quality* notion was linked to evaluation, and expanded thus the evaluation field." (op.cit:106).

Telhaug og Mediås skriver også om dette skiftet på den utdanningspolitiske dagsordenen. I boken *Grunnskolen som nasjonsbygger. Fra statspietisme til nyliberalisme* kan vi se at nodalpunktet<sup>47</sup> har endret seg fra å være solidaritet, likhet og rettferdighet til å bli kvalitet, effektivitet og konkurranse (Telhaug & Mediås 2003:300:3001). Forfatterne viser til at den politiske tenkningen i Norge i 1980-og 1990-årene frem til begynnelsen av 2000-tallet, ble sterkt påvirket av to forhold:

---

<sup>46</sup> Jamfør NPM.

<sup>47</sup> Telhaug og Mediås bruker ikke betegnelsen "nodalpunkt".

globalisering med internasjonal liberalisering og arbeidsløshet. Dette medførte et fokus på konkurransen som virkemiddel, og man så at tiden var inne for en viss privatisering, sterkere omstillingsevne i næringslivet og vilje til å vurdere offentlig som privat virksomhet etter resultatene av den. Det ble etter hvert tale om å redusere statlig eierskap, og i den politiske retorikken fikk verdier som vekst, produktivitet, effektivitet og innsats sin renessanse (ibid).

### 5.3.2. NPM, evaluering, kvalitet og den refleksive modernitet

Evaluerer er i dag utviklet seg til å bli en tidstypisk standard, hevder Dahler-Larsen og Krogstrup (2000:286). De forklarer dette ved å vise til den refleksive modernitet der tidligere tiders skapte polare begreper, som for eksempel samfunn-natur, kapitalistisk demokrati-kommunisme og fremskritt-stagnasjon, er under avvikling. I stedet for kan man ane nye problemer og utfordringer i samfunnet: Det er samfunnets egne bieffekter og koordineringsproblemer som setter dagsorden (ibid). Dahler-Larsen og Krogstrup mener institusjonaliserte verktøy innenfor organisasjon, økonomi, vitenskap og hele det sosiale felt skaper bieffekter for seg selv og for hverandre (ibid). Dette kan tolkes dit hen at samfunnets diskurser har blitt så komplekse at det er vanskelig å kontrollere den videre utviklingen. Med referanse til Beck kan man hevde at årsaken til denne kompleksiteten er at samfunnet ikke lenger har et senter med et samlet overblikk. Han skriver:

”In the old Europe, people always described the polity according to the model of the human body. The government was a supreme ruler, the head. This metaphor is definitively passé. No centre that predicts, controls and decides is available any more. The brain of society can no longer be localized; innovations and decisions on the future have not originated from the political class for some time now. On the contrary, only when an idea has become a banality does the coin drop for parties and government.” (Beck 1994:39).

Man kan kanskje tolke dette dit hen at offentlig virksomhet har, på grunn av den refleksive modernitet, mistet noe av sin styrende og kontrollerende funksjon, og at det er markedet og dets verdier og oppskrifter på god organisering som er bestemmende for dagsorden. Evalueringer kan således være med på å legitimere bruk av ulike oppskrifter



og verktøy. Dahler-Larsen og Krogstrup kan muligens undergrave dette idet de mener at man i den refleksive modernitet aldri kan ha full kontroll og treffe rasjonelle avgjørelser. De hevder at det som virkelig skjer er at viktige avgjørelser tas ”via vedvarende og ukoordinerte feedback-løkker på forskjellige nivåer i komplekse systemer.” (Dahler-Larsen & Krogstrup 2000:286).

Det *refleksive* i den refleksive modernitet må her sees på som både den ukoordinerte tilbakekastingen av samfunnets bieffekter på seg selv og et *refleksjonsarbeide* som viser til den tvil og kritikk som moderniteten retter mot seg selv. Dermed følger det at tvil og kritikk produseres av profesjoner og eksperter som setter søkelyset på problemer og utfordringer og deretter fremmer reformforslag uten å være sikre på at de i neste omgang selv blir utfordret av en liknende kritikk (op.cit:286-287). I diskursteoretiske termer kan man hevde at i den refleksive modernitet vil man aldri kunne være sikker på at meningsfikseringer vil vedvare. Det foregår hele tiden kamper mellom ulike diskurser som prøver å oppnå hegemoni.

Man kan vise til Bauman for å belyse dette ytterligere. Han reflekterer over hvordan begreper om kvalitet har blitt medier for refleksjon, tvil og omstillingsevne, men som ikke kan, under de aktuelle betingelsene, selv skape en visjon av kollektiv forpliktende karakter (Bauman 1995:79). Ut fra dette kan man kanskje identifisere en forklaring på hvorfor kvalitetsbegrepet brukes på den utdanningspolitiske dagsorden i dag, og hvordan denne bruken begrepsmessig er problematisk. Kvalitetsbegrepet kan sies å ha en rekke fordeler da det er et begrep som alle tenker på som noe positivt. Det er et begrep som gir en rekke retoriske fordeler<sup>48</sup>, og som dermed kan brukes i sammenhenger om positiv utvikling og forbedring. Begrepet forbindes med refleksjon og fremgang, uten å gi noen klare kriterier for hvordan dette skal gjennomføres. Her kan det etter min mening ane et paradoks: Når man bruker kvalitetsbegrepet hyppig på den utdanningspolitiske dagsorden, når kvalitet blir brukt i enhver sammenheng og til enhver tid, kan det miste sin refleksive verdi, det kan bli et trivielt begrep som blir liggende utenfor ens egen virkelighet. Man kan kanskje dermed hevde at begrepet blir et skapt symbol som ikke internaliseres i den enkelte fordi refleksjonen ikke lenger er til stede. Det kan i denne forbindelse muligens være belegg for å tolke at denne utviklingen

---

<sup>48</sup> Jamfør Henmo i kapittel 5.3.

særlig kan vise seg når økonomiske begreper og markedets oppskrifter knyttes til kvalitetsbegrepet.

#### 5.4. Internasjonalisering og teknologi

Internasjonalisering og teknologi kan oppfattes som to sterke krefter som har betydning for spredningen av NPM-filosofien, evalueringsideer og organisasjonsoppskrifter. Den internasjonale dimensjonen kan sies å virke både som leverandør av modeller og som generell inspirasjonskilde. Internasjonale leseundersøkelser er et eksempel på dette. Det kan se ut som om disse har en sentral rolle i utviklingen av den nasjonale skolepolitikken. Albæk, Dahler-Larsen, Foss Hansen, Krogstrup, Larsen, Lassen og Winter mener det også er mulig å se en indirekte påvirkning, i tillegg til den direkte påvirkningen i form av tall og resultater til sammenlikning, i form av bestemte måter å tenke på, for eksempel forestillingen om at alle barn i tredje klasse skal kunne lese på et gitt nivå og like godt i forskjellige land<sup>49</sup> (Albæk m.fl. 2001:263).

Teknologien kan også sies å spille en rolle i denne sammenheng. Stadig billigere og mer tilgjengelig datateknologi gjør det mulig for store og små organisasjoner å sette i verk løpende kvantitative overvåkningssystemer. Enkle monitoreringssystemer gjør det mulig å avlive myter og generere viten om kvantitative forhold som gjør det mulig å rette søkelyset mot spesifikke problemstillinger som må gjøres til gjenstand for ytterligere undersøkelse. Albæk, Dahler-Larsen, Foss Hansen, Krogstrup, Larsen, Lassen og Winter poengterer at det er viktig å merke seg svakhetene i forbindelse med monitoreringssystemer (ibid). De mener at den enkle målingen aldri kan vise hele bildet av innholdet i en offentlig organisasjons arbeid, den kan lede oppmerksomheten bort fra andre deler som kanskje er viktigere. Et annet kritisk punkt de fremhever ved slike kvantitative målingssystemer er at de kan medvirke til at selve monitoreringssystemet blir veiledende for oppgaveutførelsen fremfor oppgavens innhold og karakter (op.cit:264). Det er svært viktig at det som måles ikke fjerner oppmerksomheten vekk fra det som ikke måles. Man kan selvfølgelig ikke skylde på datateknologien i seg selv for denne utviklingen, det er måten teknologien brukes i organisatorisk gjeldende

---

<sup>49</sup> Den henvises her til PISA-undersøkelsene (Programme for International Student Assessment) som gjennomføres i OECD-landene.

tolkninger og beslutninger som er det sentrale her. Et illustrerende eksempel man kan vise til er hvordan den økende bruken av *benchmarking* som styringsredskap kan sees som en viktig faktor bak den tiltagende interessen for kvantifiserbare målinger av det offentliges prestasjoner (ibid).

I følge House kan man hevde at det ikke lenger er slik at de nasjonale regjeringene har total kontroll over alle sider ved det nasjonale liv, spesielt ikke den nasjonale økonomien (House 1990). Det internasjonale handelsmarkedet kan sies å binde landene sammen. Dette kan medføre at økonomiske beslutninger som påvirker vårt nasjonale liv, foretaks på steder langt vekk, og baserer seg på andre grunner enn de som tjener nasjonen. House mener ”markedsøkonomien bringer alle fasetter av samfunnet inn i sin sfære, og har en destabiliserende effekt på samfunnet fordi den oppløser tradisjonelle strukturer og sosiale relasjoner.” (op.cit:154). Denne internasjonaliseringen kan bli reflektert i krav om utdanning som skal produsere visse typer elever og kunnskap, slik at nasjonen kan konkurrere med suksess på det internasjonale markedet. House mener at et press som dette i mange land har resultert i en økende sentralisering av utdanningssystemet og i normer som blir fastlagt ovenfra i form av standardiserte tester og nasjonale penser (ibid). Han trekker frem USA og Storbritannia som eksempler på denne trenden. Men det er ikke på noen måte sikkert at innsnevring av fokus til bare å gjelde de sentraliserte nasjonale utdanningssystemene, vil resultere i de fordelene som tilhengerne sier, det vil si et mer effektivt, nasjonalt økonomisk system som vil konkurrere fremgangsrikt med Japan (ibid). ”Faktisk er det slik”, skriver House, ”at dokumentasjon angående nylig gjennomførte utdanningsreformer i USA, som er blitt gjennomført av politiske myndigheter uten at fagfolk er blitt konsultert, viser at utdanningssystemene i økende grad holder på å bli ineffektive og uproduktive – som i sin tur krever flere sentraliserte reformer. Med andre ord er effekten av reformene motsatte av det reformatorene tenkte.” (op.cit:154-155).

#### 5.4.1. Globaliseringens paradoks

Man kan antakelig betrakte internasjonalisering og teknologi som en trussel mot nasjonalstaten og de nasjonale staters muligheter til å styre samfunnsutviklingen etter

egne politiske mål. Karlsen skriver i *Utdanning, styring og marked: norsk politikk i et internasjonalt perspektiv* at nasjonalstaten og den nasjonale regjering blir for liten som beslutningsenhet i en global verden. Statens muligheter for å regulere den nasjonale økonomien svekkes etter hvert som globaliseringstendensene øker (Karlsen 2002:70-71). Samtidig kan man se at den statlige styringen omformes slik at den utøves etter de samme prinsipper som gjelder for styringen av det private marked<sup>50</sup>. Dette kan tolkes dit hen at staten som regulerende instans viker til fordel for en stat som i første rekke skal skape gode rammebetingelser for deltakelse i den globale konkurransen. En slik utvikling kan resultere i en statlig tilbaketrekking og forskyvning i oppmerksomhet fra velferd til effektivitet (Telhaug & Mediås 2003:286).

I denne globaliseringssammenheng kan man nok si at det er mulig å identifisere et paradoks i utdanningspolitikken: Organisasjonsoppskrifter som NPM og kvalitetsvurdering brukes som løsning på organisasjonsinterne problemer, de skal finne positive og negative sider i vårt eget utdanningssystem og gi løsninger på hvordan disse skal håndteres. Men, mister vi da muligheten til å kalle vårt nasjonale utdanningssystem *vårt eget*? Oppskriftene og løsningene stammer fra globaliseringens krefter og markedets begrepsapparat, hva er det da de finner som problemer, og ikke minst; hvilke løsningsforslag viser seg i en slik tankegang? Man skulle tro at det dermed ble mer og mer viktig å delta i en debatt om hvilke negative konsekvenser globaliseringen har, men det er vel egentlig det motsatte som skjer. Globaliseringen viser seg som en samtidsdiagnose, og den aksepteres som et faktum på tvers av tradisjonelle skillelinjer i politikken (op.cit:288).

## 5.5. NPM og utviklingen av verktøy for effektivitet og kvalitet:

### Indikatortenkning

Man kan hevde, med referanse til fremstillingen i kapittel 5.4, at globalisering og internasjonal økonomi har tegnet opp nye rammer som påvirker politikktutformingen på alle plan. Nye tiltak og reformer argumenteres med effektivitet og økonomisk vekst. Dette kan tolkes dit hen at troen på fremskrittet, særlig økonomisk vekst, står i sentrum

---

<sup>50</sup> Jamfør NPM.

for retorikken som utspiller seg i dagens politiske kontekst. Her kan man muligens spore nok et dilemma: De økende kravene til produksjon og effektivitet i kunnskapssamfunnet krever *mer* utdanning. Den tradisjonelle måten å møte slike krav på, har vært å sette inn nye økonomiske ressurser, men nå begrenses ressursbruken samtidig som det kreves *mer for hver krone*. Dette medfører at man må tenke nytt i forhold til styring av utdanningssystemet. Man ser behovet for en styring som beveger seg i retning av monitorering av systemet. Man kan spore et sterkere behov for kontroll i hele utdanningssystemet. Man kan kanskje hevde at evalueringens hovedfunksjon ikke lenger er utvikling, men kontroll. *Konkurranse* utvikler seg i denne sammenheng til å bli et av de viktigste virkemidlene. Man ser at det fokuseres på indikatorer som kan måle og sammenlikne, det innføres incentiver av diverse art, som blant annet demonstrasjonsskoler, bonusskoler, individuell belønning av ansatte og fritt skolevalg, og man argumenterer for behovet for økonomisk vekst på den utdanningspolitiske dagsordenen. Den utdanningspolitiske dagsorden må legitimere denne ”nye” kontrolleringen ved å vise til termer som globalisering, internasjonalisering, effektivitet og økonomi. Med en slik retorikk kan man hevde at skolens dannende funksjon svekkes og at kunnskapen blir vurdert ut fra mål som befinner seg utenfor skolens egen livsverden (Telhaug & Mediås 2003:322).

Indikatortenkning kan sees på som en måte å møte de nye utfordringene på. I tillegg kan denne tankegangen betraktes som en legitimeringsstrategi fra politisk hold. Her kan man si at indikatorer blir sett på som en betegnelse for hva som har betydning i forhold til læring, og dermed *kvaliteten* på skolen. Indikatorenes<sup>51</sup> funksjon er å hjelpe skolene til å oppdage viktige områder der skolene har sine sterke sider, oppdage områder som trenger forbedring og utarbeide handlingsplaner på grunnlag av funnene. Denne målingen har i mange sammenhenger blitt sett på som evaluering. Den utdanningspolitiske dagsorden kan bruke målinger som dette for å legitimere egne handlingsplaner og henvise til at det er gjennomført grundige evalueringer. Med min forståelse av evaluering kan man ikke kalle en ren indikatormåling for evaluering, da denne målingen mangler viktige dimensjoner i et evalueringsforløp. En slik måling kan kun kalles – ja, nettopp – en måling.

---

<sup>51</sup> Indikatorer blir også kalt kriterier. Internasjonalt snakkes det om *performances indicators* og *performance measurement*, som forkortes med henholdsvis PIs og PM.

Perrin er opptatt av disse problemstillingene rundt indikortenkning; performance indicators og performance measurement. Han skriver: "...the only thing new about "performance measurement" is its name. PM "reforms" have been recycled since the 1960s." (Perrin 1998:368). Videre viser han til et eksempel fra møbelindustrien i Polen: "In Poland under communism, the performance of furniture factories was measured in the tonnes of furniture shipped. As a result, Poland has the heaviest furniture on the planet." Perrin kritiserer det "nye" arbeidet med PM for ikke å lære av historien og hva som har skjedd i tidligere tilsvarende arbeid. Han skriver: "Without taking into account past experience with similar approaches to PM, the current fascination with PM runs the risk of being just another management fad which too shall pass in the course of time. Lack of acknowledgement of earlier approaches precludes the opportunity to learn from and to build upon past experiences." (ibid). Perrin understreker at det viser seg i økende grad at dagens tilnærming til PM repeterer de samme feilene som har blitt gjort i fortiden.

Mintzberg støtter Perrin i sin bok *The Rise and Fall of Strategic Planning*. Han peker på feilene ved PM-systemer i både offentlig og privat sektor. Mintzberg indikerer at disse og liknende målingsbaserte tilnærminger ikke har resultert i økt *performance* eller et fokus på resultat. Man kan heller si at de har forhindret utviklingen og implementeringen av effektive strategier (Mintzberg 1994). Perrin understreker dette ved å vise til Vietnamkrigen hvor antall døde fiender var en indikator på fremgang i krigen. Han spør: "How many times do we have to come back to this one until we give up?" (Perrin 1998:368).

I et oranasjonsteoretisk perspektiv kan man henvise til Røvik for å underbygge Perrins påstand om at PM-reformer ikke er noe nytt, men at *oppskriftene* har blitt brukt siden 1960-tallet. Perrin mener at disse oppskriftene kan se nye og innovative ut, men sannheten er at de kun er gamle oppskrifter i ny "språkdrakt" (ibid). Røvik skriver i forbindelse med dette om utfordringen som ligger i å fange inn og forklare det *dynamiske, syklusaktige* aspektet ved moderne institusjonaliserte organisasjonsoppskrifter<sup>52</sup>, nemlig det at de oppstår og spres for så å tape seg og bli

---

<sup>52</sup> Røvik definerer en institusjonalisert organisasjonsoppskrift som "en legitimert oppskrift på hvordan man bør utforme utsnitt eller elementer av en organisasjon. Det er en oppskrift som fenger og som har fått en forbilledlig status for flere organisasjoner." (Røvik 1998:13).

fortrengt av nye oppskrifter (Røvik 1998:19). Utfordringen er i følge Røvik å vinne innsikt i ”så vel hvordan og hvorfor oppskrifter *institusjonaliseres*, det vil si blir til populære og gradvis mer tatt-for-gitt-forbilder, som at de etter tid gjerne *av-institusjonaliseres*, det vil si taper sin forbildestatus i forhold til en gruppe organisasjoner” (ibid). Videre hevder Røvik, at ved å utvide perspektivet til å for eksempel omfatte noen tiår, er det kanskje i tillegg nødvendig med begrepet *re-institusjonalisering* for å få tak i dynamikken. Re-institusjonalisering viser til at oppskrifter som har vært populære og deretter har tapt sin popularitet, senere kan dukke opp, gjerne i ny språkdrakt, og igjen få status som en forbilledlig måte å organisere på (ibid).

Perrin diskuterer ikke bare det man kanskje kan betegne som en *historieløshet*, det vil si det kritiske ved å falle for ”nye” og populære oppskrifter som viser seg å være ”gamle” og feilslåtte. Han går også grundigere inn i problemstillingene som ligger i *bruken* av PIs og spør: “Are the limitations of performance measurement due to its faulty use – e.g. poor performance indicators (PIs), or a result of more basic flaws in its logic?” (Perrin 1998:370). Perrin diskuterer problemstillingen i forhold til åtte problemer og begrensninger (op.cit:370-375):

1. Forskjellige tolkninger av samme term og konsept: PIs blir brukt og fortolket på forskjellige måter selv om de er nøye beskrevet. Dermed blir sammenlikningsgrunnlaget mellom forskjellige kontekster og personell svekket (selv om man prøver å fremstille dataene i rene mål og tall). Hvis personalet føler at fremtiden til deres program, eller kanskje til og med jobben deres, kan bli bestemt av utfallet av en indikatoremåling, er det naturlig at de tolker indikatorene i sin favør.
2. Målforskyvning/-endring: Når indikatorene utvikler seg til å bli *målet*, vil de resultere i en ”målforskyvning” som resulterer i feil fokus, på feil aktiviteter og oppmuntret til en praksis der målet er å skåre høyt i målinger, uten at det dermed resulterer i et godt resultat/læring.
3. Bruk av meningsløse og irrelevante mål/standarder: PIs kan være irrelevante, selv om de er eksakte. Det grunnleggende i en PM er å redusere et kompleks program/system til få og enkle indikatorer. Dette strider mot teorier om hvordan store system fungerer (systemteori), teorier som ser at helheten er større enn summen av delene. PIs ignorerer den kontekstuelle kompleksiteten som ligger i

sosiale fenomen, en kompleksitet som inneholder mange gjensidig påvirkende faktorer som ikke kan forstås alene, eller i et bestemt antall indikatorer.

4. Kostnadsbesparelse versus kostnadsendring: PIs brukes som individuelle isolerte program, fjernet fra konteksten. Konsekvensen av dette er at ”resultatene” kan representere en kostnadsendring heller enn en sann kostnadsbesparelse.
5. Misvisende aggregering av indikatorene skjuler kritiske forskjeller mellom ulike undergrupper: PIs skjuler forskjellene som ligger i ulike undergrupper.
6. Begrensninger som viser seg i målbaserte tilnærminger til evaluering: PM systemer deler de veldokumenterte begrensningene som vi finner i andre målbaserte tilnærminger til evaluering.
7. Ubrukelig for avgjørelser og forklaring av ressursbruk: Det vanligste argumentet *for* PM er å sørge for bedre informasjonsgrunnlag før valg av tiltak og budsjett fattes. Men PM er i seg selv ubrukelig til dette. Dette er fordi PM ikke spør de grunnleggende hvordan- og hvorforsspørsmålene.
8. Skaper mindre resultatfokus: Det ironiske, og triste, er at PM systemer ikke resulterer i et større fokus på resultat, innovasjon og forbedring.

Perrin sikter ikke her til at man må forkaste all bruk av PM/PI, men det er en del faktorer som må tas i betraktning når man vurderer å bruke denne tilnærmingen (op.cit:375). Først er det viktig å innse at bruk av PIs passer best for planlegging og monitorering, *ikke* som evaluering. De kan spille en rolle i arbeidet med å avdekke fokusområder, det vil si de kan hjelpe til å sette fokus på visse områder som trenger videre oppmerksomhet og evaluering. Videre må man forstå at enhver evalueringsmetode har sine begrensninger som bare kan overkommes ved å bruke flere metoder samtidig. Det er vanligvis nødvendig å sammenlikne kvantitative data fra PIs med kvalitative informasjonsdata. I tillegg er det grunnleggende å være oppmerksom på at PIs kan være hensiktsmessig å bruke i noen tilfeller, mens i andre kan de være totalt ubrukelig. For at PM skal kunne brukes, må programmet ha tilstrekkelige ressurser, inkludert metodisk ekspertise i utviklingen av PIs, interessentene må involveres i utformingen og bruken av målingsinstrumentene (indikatorene) og det må gis tid til å utvikle, teste, rense, revidere og oppdatere målingsinstrumentene. Perrin kan forstås som en skeptiker til bruk av PM/PI, noe som kan underbygges med følgende sitat: ”Unfortunately, many of these conditions are unlikely to be met in the real world, where



political realities frequently impose performance measures in a top-down manner, where accountability for results is required in a unrealistically short time frames, and where qualifications accompanying tables or graphs displaying performance indicators are routinely ignored.” (op.cit:377). Winston fremhever dette videre idet han leter etter fordelene ved PM: ”To hear of the benefits of PM a thousand times still leaves the requirement to see them at least once.” (Winston 1999:98).

Det kan se ut som om forkjempere for PM fokuserer på hvordan PM er planlagt å fungere. Perrin mener han selv, og andre som er kritiske til PM, prøver å indikere hva som egentlig skjer i den virkelige verden, hvor organisatoriske og politiske pressituasjoner stadig misbruker tilnærmingen (Perrin 1999:102). Perrin mener å se stor grad av naivitet når det gjelder disse realitetene, realiteter som oftest resulterer i feil bruk av resultater. Her er det relevant å diskutere hva som er oppgavene til evaluator<sup>53</sup>. Chelimsky sier: ”our most important task and the best justification for our work as evaluators is our capability and willingness to question conventional wisdom.” (op.cit:106). Og Perrin viser videre til Kuhn for å utdype: “As Kuhn pointed out, the tendency is to reject evidence that does not fit the prevailing paradigm, rather than the other way around. Logical positivism, along with reductionist schemes such as PM, dies hard.” (ibid).

Bernstein mener det ikke bare er interessentene som feiltolker og misbruker PM (Bernstein 1999:90). Det er nødvendig med en forståelse av forholdet mellom evaluering og PM. Evalueringsprofesjonen har verken teoretisk eller praktisk forsket på relasjonen mellom PM/PIs og evaluering. Han mener en slik gransking av PM og evaluering kan utvikle den teoretiske og praktiske forskjellen. For det første er det slik at ikke-evaluators ofte er ansvarlig for å rapportere informasjonen, dermed oppstår det en inadekvat oppmerksomhet til programmodellering og validitetsspørsmål. For det andre har det vært lite oppmerksomhet omkring bruken av PM, om hvordan resultatene fremstilles, hvilken nytteverdi målingen har og mistolkning og misbruk av resultater. Bernstein er en evalueringsteoretiker som er positiv til at evaluator må gå inn i debatten etter en evaluering og si sin mening. Det er evaluators oppgave å formulere verdimesig utsagn om evalueringsresultatene, og ikke lide av verdifobi. I tillegg understreker han

---

<sup>53</sup> Evaluators oppgaver diskuteres videre i kapittel 6.1. Der diskuteres oppgavene i forhold til en demokratisk debatt.

hvor viktig det er å nå ut til omverdenen for å assistere andres arbeid. Og, kanskje aller viktigst, vi må lære av andre og arbeid som er gjennomført tidligere. Bernstein refererer til Santayana: "Those who do not remember the past are condemned to repeat it." og Napoleon: "History is the version of past events that people have decided to agree upon." for å understreke poenget (op.cit:92).

#### 5.5.1. Standardisering

Dahler-Larsen og Krogstrup reflekterer i artikkelen *Evalueringers konstitutive virkninger* over standarder og bruk av standardiseringer, PIs, i organisasjoner. En standard kan teoretisk defineres som "en regel som sier noe om hva som bør gjøres i bestemte situasjoner" (Dahler-Larsen & Krogstrup 2001a:237). De kan oppfattes som råd til andre om hva de kan gjøre, men de kan også bli så institusjonalisert at de blir objektivert og får karakter av normer. Dermed kan de bli til direktiver hvis de vedtas som en del av en organisatorisk praksis, eventuelt med tilknyttede sanksjoner, hvilket ofte er tilfellet med evalueringer, i følge Dahler-Larsen og Krogstrup.

I artikkelen formuleres det fire sentrale premisser for standardiseringer, eller PIs (ibid). Det første er at det er mulig å *avgrense prosesser*. Bak denne tankegangen ligger det en forståelse av at prosesser har eiere og at det er disse eierne som også har kontroll og ansvar for det som skjer. Men virkeligheten er langt mer kompleks enn som så. For eksempel har media en vesentlig innflytelse på etableringen av nye tiltak og dermed prosesser i den offentlige sektor. Det andre premisset bygger på forestillingen om at *organisasjoner kan styres* og at organisasjoner betraktes som maskiner hvor enkelte komponenter (for eksempel medarbeidere) kan skiftes ut med tilsvarende komponenter uten at dette får særlige konsekvenser. Ledelsens oppgave blir da å styre mennesker og teknikker slik at målene oppfylles. Mange organisasjonsteoretikere har vist at slik vil man ideelt ha det, men at virkeligheten er en helt annen (ibid). Et tredje premiss, som gjelder noen typer standardiserte evalueringer, er forestillingen om at organisasjoner skal være *responsive i forhold til kundenes behov*. Dette kan være problematisk for offentlige organisasjoner da de har mange andre hensyn å ta enn bare å imøtekomme kundenes behov. Det fjerde premisset som ligger til grunn for å måle innsatsen ved

hjelp av PIs, er at det er mulig å snevre inn kundenes behov til *mål som er målbare*, slik at det kan dokumenteres om tjenesten som leveres utfyller disse behovene.

Det problematiske med en slik standardisering er at den kan virke mot sin hensikt, det vil si at man blir flink til å følge standarden i stedet for å forholde seg refleksivt, kritisk og faglig til oppgavene som skal utføres. Da kan man hevde at evalueringskriteriene blir verdier i seg selv. Dahler-Larsen og Krogstrup mener i denne sammenheng at evalueringen setter sine egne tidsmessige rammer og er med på å forme aktørene (op.cit:239-240). De belyser dette ved å si: "Evalueringer beskriver ikke bare virkeligheten. De former den." (op.cit:243). Hvor mye og hvordan det skjer, er avhengig av de sosiale, organisatoriske og politiske kontekster. Hvilken rasjonalitet preger disse kontekster?

Man kan her vise til Bauman for å belyse dette spørsmålet. Bauman betegner handlinger som styres av reglenes fornuft, for eksempel handlinger som styres av PIs, som *prosedyrerasjonalitet* (Bauman 1995). Man kjenner igjen denne rasjonalitetstypen først og fremst ved en sterk lojalitet overfor organisasjonen og en villighet til å utføre sine plikter. Bauman hevder at en konsekvens av prosedyremessig rasjonalitet kan være at den handlende aktør neglisjerer sine moralske og følelsesmessige overveielser og dermed fraskriver seg et personlig ansvar. Dette kan man gjøre ved å henvise til at man som medlem av organisasjonen må følge de vedtatte prosedyrene. De kortsiktige konsekvensene for mennesker som utsettes for standardisering begrunnet i prosedyremessig rasjonalitet, kan være at moralske spørsmål tas av dagsorden, og at de handlende individers mulighet for moralsk bedømmelse undergraves; de fraskriver seg det moralske ansvaret og ignorerer sine egne handlingers konsekvenser (op.cit:259). Dahler-Larsen og Krogstrup mener at den beste måten å motvirke disse uheldige konsekvensene, er å se på evalueringer som verdiinnlegg heller enn som nøytrale, tekniske målinger (Dahler-Larsen & Krogstrup 2001a:243-244).

### 5.5.2. Benchmarking som strategi

”Å sammenlikne er å forbedre”. Dette kunne kanskje vært et slagord for benchmarking, en metode som er utviklet for næringslivet, men som ser ut til å bli brukt mer og mer av det offentlige. Benchmarking innebærer at man systematisk sammenlikner seg med konkurrenter (og andre) på en rekke nøkkelindikatorer, og gjennom det ser klart for seg forskjeller og forbedringsmuligheter (Røvik 1998:257). Røvik refererer til Robert Champ<sup>54</sup> som den som har utviklet benchmarking til en organisasjonsoppskrift. Benchmarkingprosessen blir gjerne fremstilt i fem steg (op.cit:258-259). Det første steget er å *velge ut hva som skal sammenliknes*. Man må bestemme seg for hvilken del av virksomheten som trenger en forbedring, og som dermed trenger å sammenliknes med andre. Røvik hevder at det i de fleste fremstillinger av konseptet legges vekt på at det er prosesser (arbeidsmetoder, rutiner, beslutningsfatting osv.) og ikke produkter som bør sammenliknes. Det andre steget i prosessen er å *bestemme hvem man skal sammenlikne seg med*. Det er viktig å velge ut en god sammenlikning. Videre skal man *innhente informasjon* som har relevans i forhold til de variabler man vil sammenlikne. Denne informasjonen bør komme fra åpne kilder. Det fjerde steget er *analyse*. Det er her den egentlige benchmarkingen skjer; man sammenlikner den innhentede informasjonen med tilsvarende prosesser i egen organisasjon. Det femte, og siste, steget beskrives som *tiltak for forbedring*. Analysen som gjennomføres skal ende opp med tiltak som kan hjelpe virksomheten til å forbedre seg.

I denne sammenlikningen kan man hevde at det finnes et konkurranseelement. Dette behøver ikke å bli uttalt offisielt, men kan ligge i bunn som et strategisk grep. Man kan kanskje si at det er lettere å argumentere *for* en slik strategi ved å vise til at sammenlikningen er til for å finne styrker og svakheter i eget system, og ikke for å måle seg med tilsvarende system. ”Det er i sammenlikning med andre systemer du kan se ditt eget” kan oppfattes som et argument *for* benchmarking som kan uttales offisielt. Men konkurranseelementet kan også sies å være sentralt her. Slike vurderinger kan utvikle en holdning som likner på idrettsdeltakere som vil på toppen av ”pallen”. Media kan dekke ”sportsevenen” og diskutere Norges innsats i forhold til de land vi liker å sammenlikne

---

<sup>54</sup> Champ, R.C. 1989. *Benchmarking. The Search for Industry Best Practices that lead to Superior Performance*. Milwaukee, WI: ASQC Quality Press.

oss med. Dermed kan man spore problemer i eget system som man er nødt til å ”trene” mer på. Hvem er det som skaper disse problemene? Er det *ekte* organisasjonsinterne problemer?

Symbolperspektivet er, etter min mening, fremtredende her. Hvis benchmarking brukes som verktøy for planlegging på den utdanningspolitiske dagsorden, argumenteres det antakeligvis ut fra at man i en slik måling ser hvilke styrker og svakheter som er å finne i ens eget utdanningssystem. Målingen kan gis rasjonalitet ved at den gir klare tall på hvor man plasserer seg på en skala i forhold til andre land. Målingen, og verktøyene som brukes i etterkant av målingen, har mulighet til å oppnå legitimitet hvis den utdanningspolitiske dagsorden plukker ut de resultatene som legitimerer egne valg og handlinger. Man kan vise til internasjonal forskning og håpe at dette holder for å oppnå legitimitet blant skolefolk og i offentligheten. Dette kan forstås som et retorisk grep som kan være skadelig både for skolens utvikling og offentlighetens forståelse av skolehverdagen og dens utfordringer. Man bør oppfatte disse resultatene kun som data som viser en avgrenset del av den virkelighet som befinner seg i skolen, og ikke som noe som viser et helhetlig bilde. Hvis resultatene kun blir brukt beskrivende og dermed kan bli forvandlet fra symboler, som skapes utenfor eget system, til rene verktøy for hele det nasjonale utdanningssystemet, kan det oppstå uheldige konsekvenser. Problem kan oppstå hvis ”evalueringen”, eller kanskje man heller bør si sammenlikningen, mangler en programteori som på forhånd er bestemt, altså at det ikke finnes kriterier for hvordan man skal tolke og verdsette resultatene. Stefan Hopmann understreker dette ved å henvise til PISA-undersøkelsen. Han hevder at denne undersøkelsen ikke kan si noe om skolens kvalitet: ”Det finnes jo ikke noe felles målestokk for hva som er god eller dårlig leseferdighet, eller mye og lite bråk mellom de forskjellige landene. Så hvordan skulle man da kunne sammenlikne dem?” (Telhaug & Mediås 2003:327).

## 5.6. NPM og modernisering

*Modernisering* er i følge Hood en tvetydig idé, noe som muligens kan forklare den sterke appellen ordet har (Vanebo m.fl. 2001:15). Han mener at modernisering i forbindelse med offentlig administrasjon og ledelse har minst fire betydninger:

1. Modernisering kan identifiseres som en klar og tydelig bevegelse fra de tradisjonelle måter å organisere og utføre offentlige tjenester på til dagens moderne måter og metoder.
2. Modernisering er en uunngåelig og irreversibel endring og fornyelse av offentlig sektor.
3. Modernisering er en endring som vil lede til konvergens, og de samme moderne måter for styring oppstår overalt (jamfør symbolperspektivet).
4. Modernisering er en endring som i vidt omfang er fordelaktig og derfor bør ønskes velkommen.

NPM som samlekonsept, og store organisasjonsoppskrifter ellers, kan sies å være modernistiske i alle disse fire betydningene. Hood mener blant annet at historien om OECDs rolle i forhold til rike demokratier tenderer mot å være modernistisk i forhold til disse fire betydningene (ibid). Han sier at påstanden om modernisering i betydningen *bevegelse* er vanlig hos praktikere og akademikere som studerer endringer i offentlig service- og tjenesteproduksjon. Påstanden om uunngåelighet ligger implisitt i de vidt spredte påstandene om et paradigmeskifte, det vil si at det er mer enn en tilfeldig motebølge som oppstår i styringen av offentlig sektor. Den tredje påstanden finnes hyppig i det som skrives om offentlig ledelse og administrasjon. Denne oppfatningen har blitt berømt med Osborne og Gaebler påstand om et nytt, globalt ”entreprenøraktig” paradigme for offentlig ledelse som erstatter avleggse, regelbundne former for organisasjonsdesign i offentlig sektor. Påstanden om modernisering av offentlig ledelse i den fjerde betydningen ligger implisitt i skrifterne til de moderne organisasjonsguruene, representert ved OECD, Verdensbanken og Osborne og Gaebler (op.cit:16).

Hood stiller spørsmålstegn ved alle disse påstandene og vedtatte sannhetene. Han mener denne visjonen om ”en gang for alle moderniseringsendring” som sveiper gjennom offentlig ledelse, overalt har appell (ibid). Det kan kanskje stilles et spørsmål om denne visjonen samsvarer med et vel etablert behov for en stor visjon om en fremtid som ikke bare er mer av det samme. Man kan si det i hvert fall her handler om de behovene som spesielle aktører har for å legitimere sin styrende og påvirkende rolle i samfunnet. Men Hood mener at man ved nærmere undersøkelse kan se ”at disse visjonære moderniseringspåstandene blir mer utfordrende fordi de mangler historisk perspektiv,

synes å være tvetydig og fordi mye avhenger av hvilken betydning av ordet modernisering vi snakker om. Modernisering har blitt credoet. NPM er løsningen.” (ibid).

#### 5.6.1. En rasjonell og kontrollerbar verden?

Kvalitetsvurdering kan sees på som et forsøk på å hankses med store organisatoriske påvirkninger som krav til effektivitet og økonomisk rasjonalisering, modernisering og et sterkt fokus på evalueringens funksjon. I denne sammenheng viser det seg et bilde av det rasjonelle, man har behov for klare og håndterlige verktøy som kan hjelpe en i et kompleks samfunn. Spenningen mellom det komplekse, uforutsigbare og uforståelige, og troen på rasjonelle verktøy for å løse dagens problemer, kan tolkes som et sentralt element i forklaringen på hvorfor kvalitetsvurdering har utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden. Giddens spør i denne sammenheng: ”Hvorfor har ”fornuftens” allmenne utbredelse ikke skapt en verden som vi kan forutsi eller kontrollere?” (Giddens 1997:110). Svaret er ikke enkelt. Det finnes en rekke faktorer som gjør seg gjeldende, i følge Giddens. Han nevner faktorer som designfeil, operatørfeil, utilsiktede konsekvenser og den sosiale vitens refleksivitet<sup>55</sup> eller sirkularitet, i tillegg til differensiert makt og verdienes rolle. Han skriver blant annet:

”Utilsiktede konsekvenser vi stadig forekomme. Dette skyldes den sosiale vitens sirkularitet, som i første omgang snarere påvirker den sosiale verden enn den naturlige verden. Under modernitetens betingelser vil den sosiale verden aldri kunne utgjøre et stabilt miljø på grunn av tilførsel av ny viten om dens natur og dens måte å fungere på. Ny viten (begreper, teorier, funn) gjør ikke bare den sosiale verden mer gjennomskuelig, men endrer også dens natur og sender den av gårde i nye retninger. ... For selv om kunnskap om den naturlige verden ikke påvirker verden direkte, inkorporerer den sosiale vitens sirkularitet elementer av naturen via den abstrakte systemers teknologiske komponenter. Dette er grunnen til at vi ikke kan bemerke oss ”historien” og uten videre dreie den i ønsket retning. ... Verden er ”én” i visse henseender, men radikalt splittet av maktforskjeller i mange andre. Og et av de mest karakteristiske trekk ved

---

<sup>55</sup> Jamfør ”den refleksive modernitet”, kapittel 5.3.2.

moderniteten er oppdagelsen at utviklingen av empirisk viten ikke i og av seg selv setter oss i stand til å velge mellom ulike verdsett.” (op.cit:110-111)<sup>56</sup>.

I Giddens sitat viser det seg en usikkerhet i dagens samfunn. Hvis man overfører dette til den utdanningspolitiske dagsorden kan det bety at den må jobbe for å gi rasjonalitet og prøve å oppnå legitimitet til noe som er usikkert, noe det ikke finnes klare svar på. Man må jobbe for å identifisere seg som rasjonelle politiske aktører for å oppnå legitimitet til politiske handlinger. Utfordringen ligger da i å skape en overbevisning om at deres politiske valg og handlinger er de riktige i en usikker hverdag og fremtid. De må videre undergrave denne usikkerheten og fremstå med valg og handlinger som kan virke rasjonelle og troverdige. Det største problemet er kanskje at det er vanskelig å si noe om de langsiktige konsekvensene av et fokus på for eksempel et kvalitetsvurderingssystem. Dette problemet vil derfor muligens oversees til fordel for de kortsiktige og rasjonelle ”her og nå”-fordelene.

## 5.7. Behov for en ny identitet?

Kapittel 4 tegnet opp et bilde av noen endringer som har foregått internt på den utdanningspolitiske dagsorden. I dette kapittelet har jeg forsøkt å skissere hvilke tendenser i samfunnet som er med på å påvirke denne dagsorden og dens fokusering på kvalitetsvurdering. I tillegg kan fremstillingen tolkes som et forsøk på å belyse hvilken argumentasjon som ligger til grunn *for* kvalitetsvurdering. Fremstillingen viser til den ”nye” NPM-filosofien som grunnlag for den påvirkningen som kommer fra samfunn og marked til offentlig virksomhet. Sett fra et symbolperspektiv er NPM-filosofien en ekstern driver, med mange forskjellige elementer, som i neste omgang fremstår som oppskrifter som reiser i tid og rom, og som skaper problemer og finner svar på tidens utfordringer på den utdanningspolitiske dagsorden. Spesielt kan man kanskje hevde at noen elementer innenfor denne filosofien gjør seg gjeldende på den utdanningspolitiske

---

<sup>56</sup> Giddens modernitetsbegrep trenger en forklaring. Giddens reserverer seg fra å bruke begreper som postmodernitet, postmodernisme, postindustrielt samfunn, postkapitalisme osv. Han sier: ”Snarere enn å tre inn i en periode av postmodernitet, beveger vi oss inn i en periode hvor konsekvensene av moderniteten blir mer radikale og mer universelle enn tidligere. Jeg vil påstå at vi hinsides moderniteten kan ane konturene av en ny og annerledes orden som er ”postmoderne”; men dette er noe helt annet enn det som for øyeblikket av mange blir kalt postmoderne” (Giddens 1997:12).



dagsorden, og dermed etablerer seg som momenter: Økonomiske begreper, kvalitet og effektivitet, evaluering, indikortenkning og måling, sammenlikning, konkurranse, mål- og resultatstyring, forbrukerperspektiv og modernisering.

Man kan oppfatte denne filosofien som problematisk i forhold til skolens tradisjonelle verdier og identitet. Har disse endringene skapt et behov for en ny type styring, en ny utdanningspolitikk? Kan man ane en bevegelse mot en markedsstyrt utdanningsfilosofi?

Man kan kanskje øyne en påbegynnende legitimitetskrise. *Er fokuseringen på kvalitetsvurdering et grep for å skape en ny identitet og oppnå legitimitet på den utdanningspolitiske dagsorden?* Kan man hevde at det i NPM-filosofien, og da særlig i *kvalitet og evaluering*, ligger en iboende retorikk som matcher en ny og passende identitet?

#### 5.7.1. Identitet, ritual og/eller symbolsk fasade?

Man kan hevde at en organisasjons identitet handler om meningskonstruksjon og er et sentralt element i institusjonaliseringsprosesser. I den refleksive modernitet kan man si at organisasjoner har problemer med å legitimere sine beslutninger og handlinger og fremstå med en identitet som fremmer et rasjonelt bilde og oppnår legitimitet.

Teknologi, organisasjon og statlige byråkratier kan ikke lenger automatisk hvile på modernitetens store fortellinger om *fornuft og fremskritt* (Dahler-Larsen 2001a:129).

Samtidig som det blir mer og mer innlysende hvordan organisasjoner skal legitimere seg, øker presset på dem for å gjøre det, hevder Dahler-Larsen. Han mener at omdreningspunktet for meningskonstruksjonen i organisasjoner har endret seg fra å henvise til modernitetens store fortellinger til å omhandle organisasjonens identitet.

Carol Axtell Rays analyse av organisasjonskultur kan være et eksempel på dette (op.cit:130). Analysen peker på et bilde av at organisasjonskulturen prøver å etterlikne det vi tidligere kjente som kultur på samfunnsnivå, men man har nå bare erstattet organisasjon med samfunn. Dette kan forstås som en forskyvning av mange forestillinger i den samfunnsskapte virkelighet. Med henvisning til Rays analyse kan man da muligens ane at det i dag er nye og andre orienteringer og prioriteringer som blir gjeldende, og som organisasjoner prøver å forankre i sin egen identitet.

Med referanse til den refleksive modernitet kan man si at organisasjoner i dag skal handle og rettferdiggjøre seg med referanse til en troverdig identitet. I en kompleks og usikker verden kan organisatoriske aktiviteter, som for eksempel planlegging, bekrefte organisasjoners symboler, noe som kan gi en oppfatning av orientering, tro, håp og identitet. På samme måte har evaluering en rekke symbolske oppgaver i forhold til en organisasjons identitet. Evaluering kan hjelpe en organisasjon med å gjenkjenne seg selv via en ytre anerkjennelse av at det handler om "en acceptabel aktør, der anvender tidens accepterte prosedurer og tankefigurer til å løse accepterte problemer på accepterte måter" (op.cit:131). Dermed blir evalueringer en karakteristikk på legitim ledelse gjennom datainnsamling. Evalueringer kan derfor utvikle seg til å bli en identitetsskapende ritualisert begivenhet i en organisasjon som både kan være til hjelp for å dempe kritikk og stabilisere organisasjonens forståelse av seg selv i verden (op.cit:132).

Evaluering viser dermed omverdenen at man er en seriøs aktør i det organisatoriske samfunn. Evaluering er påkrevd, og alle vet at de må evaluere. Men, skaper evaluering forandring og utvikling, eller er det kun en symbolsk fasade?

Hvorfor har kvalitetsvurdering utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden? Med referanse til nyinstitusjonell teori og den refleksive modernitet kan man belyse dette spørsmålet ved å vise til at kvalitetsvurderingsfokusert både er et tegn på å skape identitet og rutiner, og er en symbolsk fasade. Man vil vise seg som en seriøs aktør i omverdenen for å oppnå legitimitet, man vil skape nye rutiner for hva som er god organisering og man vil fremstå med en organisatorisk felles identitet for å vise at "dette er oss".

Man kan hevde, med referanse til Dahler-Larsen og Pollit (op.cit:84) at NPM har vært medvirkende til å spre en norm om at offentlige organisasjoner skal evalueres. Pollit kommer i denne sammenheng med en interessant kommentar: Bemerkelsesverdig nok har selve filosofien NPM imidlertid selv vært friholdt for evaluering (ibid). Dermed kan man kanskje tolke NPM dit hen at filosofien har utviklet seg til det som James March kaller *protected discourse*, at den blir oppfattet som den eneste riktige måten å organisere offentlig virksomhet på, og derfor ikke trenger å stille spørsmål ved

(op.cit:17). Hvis man ser på NPM på denne måten, kan man hevde at filosofien har vunnet hegemoni i det diskursive feltet og slutte seg til Hoods påstand om modernisering og NPM: "Modernisering har blitt credoet. NPM er løsningen" (Vanebo m.fl. 2001:16).

## 6. HVORFOR KVALITETSVURDERING?

Hvorfor kvalitetsvurdering har utviklet seg til å bli et sentralt element på den utdanningspolitiske dagsorden er ikke et enkelt spørsmål som kan besvares med ett enkelt svar. Kvalitetsvurdering bør forstås som et kompleks fenomen, og det har hele tiden, gjennom denne oppgaven, vært et mål for meg å vise denne kompleksiteten og utvide forståelsen for den. Kvalitetsvurdering kan oppfattes som et system som skal løse skolens utfordringer i dag og i fremtiden. Det å innføre et nasjonalt *kvalitetsvurderingssystem* kan se ut til å være et riktig grep å ta i dagens samfunn, men det er viktig at man i den sammenheng gjør seg flere grunnleggende historiske, teoretiske og samfunnsmessige refleksjoner slik at ikke systemet bare blir en døgnflue, eller at det skaper utilsiktede, kanskje negative, konsekvenser i fremtiden. Mitt grunnlag for å gi svar på problemstillingen om *hvorfor kvalitetsvurdering* ligger i utviklingslinjene jeg har vist til i kapittel 4 og i de aktuelle samfunnstendensene som jeg har referert til og diskutert i kapittel 5.

Underveis i kapitlene har jeg presentert og diskutert utvalgte diskurser som kan sies å være av betydning for fokuset på kvalitetsvurdering på den utdanningspolitiske dagsorden, og som kan sees på som en argumentasjon *for* kvalitetsvurdering. Diskursene ser ut til å påvirke den utdanningspolitiske dagsorden; de kjemper for å vinne plass og oppnå hegemoni.

Historisk kan man se en utvikling mot profesjonalisering av evalueringsfeltet og vektlegging av dokumentasjon i hele skolesystemet. Utviklingen viser en linje fra testing av elevprestasjoner til evaluering og kontroll av hele systemet. Man kan se den økende vektleggingen av arbeidet med nasjonale vurderingssystemer, i tillegg til påvirkning fra internasjonale forskere og trender.

Utviklingslinjene viser også hvordan evaluering har hatt betydning for styringen av utdanningssystemet. Spørsmålet om sentralisering eller desentralisering kan oppfattes som sentralt her. Kvalitetsvurdering kan sees på som et desentraliseringstiltak der det lokale tiltenkes større handlingsfrihet og skolene/skoleeier er ansvarlige for utviklingen

og kvaliteten på skolene<sup>57</sup>. Bak denne desentraliseringen kan man hevde at det likevel skjuler seg en sentralisering; nasjonale styresmakter kan argumentere for et sterkere behov for kontroll innenfor noen områder i et desentralisert utdanningssystem. Denne kontrollen kan ivaretas med nasjonale kvalitetsvurderingssystemer der man kontinuerlig må rapportere inn resultater og dokumentere sin virksomhet. I en slik ”pakke” kan det fra nasjonalt hold også opprettes forskjellige incentivsystemer knyttet til kvalitetsvurderingen. Kvalitetsvurderingssystemer med dokumentasjonsfokus og incentivsystemer kan peke i retning av resultatstyring.

Kapittel 5, som handlet om NPM og tendenser som i dag kan sies å gjøre seg gjeldende på den utdanningspolitiske dagsorden, kan som sagt også være med på å belyse problemstillingen. Kapitlet skisserer et bilde av NPM som den nye filosofien i offentlig virksomhet og utviklingen mot *kvalitetsbølgen* og *evalueringsbølgen*, noe som kan vise hvilken eksplosjon som har skjedd i forhold til bruken av begrepet kvalitet og fenomenet evaluering. NPM kan sies å være et sentralt element å vise til når man snakker om *hvorfor kvalitetsvurdering*, og *hvilken argumentasjon* som legges til grunn for denne. Nye verdier har vist seg i det offentlige og blitt et forbilde for samfunnet; markedets verdier med økt produksjon, økt avkastning og økt konsum har blitt et forbilde for samfunnslivet som sådan (Vetlesen 1998). Det er rimelig å anta at denne utviklingen har stor betydning for hvilke styringsfaktorer som gjør seg gjeldende på den utdanningspolitiske dagsorden. I lys av dette kan man hevde at dyktighet og effektivitet kreves i større og større grad, også i skolesystemet. Videre kan man da muligens si at det viser seg et behov for administrative reformer hvor nøkkelordet er NPM (Rhyn 2003:88).

NPM kan tolkes som et konsept som stammer fra det private marked og som dermed bringer med seg næringslivets begreper, argumenter og verdier inn i det offentlige. Oppsummerende kan man si at NPM handler om markedsorientering, bruker-/kundeorientering, kontraktstyring, mål- og resultatorientering, brukerbetaling, profesjonell ledelse, desentralisering av oppgaver og konkurranse. I tilknytning til dette er det mulig å si at evaluering og kontroll nedover i systemet står sentralt i denne

---

<sup>57</sup> Accountability-begrepet kan brukes i tilknytning til desentralisering. Dahler-Larsen (2001b) poengterer dette da han hevder at det ofte knyttes et krav om ”accountability” til desentralisering av ansvar for gjennomføring av oppgaver. Accountability kan beskrives som ”ansvarlighet for arbeidsprosesser og resultater” (op.cit:79).

filosofien. Dahler-Larsen påpeker evalueringsfokuset ved å hevde at evaluering, eller noe som likner evaluering, spiller en avgjørende rolle i tidstypiske nøkkelord som kvalitetssikring, målstyring og resultatstyring (Dahler-Larsen 2001a:10). Her kan det være hensiktsmessig å legge til fenomenet kvalitetsvurdering, et fenomen som man muligens kan si prøver å favne inn alle disse nye utfordringene i et helhetlig system.

*Hvorfor kvalitetsvurdering* kan videre belyses med det som i nyinstitusjonell teori kalles *den refleksive modernitet*. Dette kan karakteriseres som en samfunnstilstand der det er samfunnets egne bieffekter og koordineringsproblemer som setter dagsorden. Dahler-Larsen og Krogstrup mener dette skjer på grunn av at samfunnet blir mer og mer komplekst (Dahler-Larsen & Krogstrup 2000:286). I den forbindelse kan kvalitetsvurdering oppfattes som et forsøk på å hankses med denne kompleksiteten; man prøver å håndtere de mange og komplekse utfordringene som skolen står overfor i dag ved å innføre et system som kan virke som et stabilt og rasjonelt verktøy. Dette kan sees på som et forsøk på å skape et bilde av overblikk og kontroll. Innføringen av et kvalitetsvurderingsvurderingssystem kan dermed tolkes som et forsøk på å gi et bilde av at den politiske dagsorden har forståelse for, og kontroll, over dagens situasjon og fatter de beste valgene for fremtidens skole.

Internasjonalisering og teknologi kan også oppfattes som betydningsfulle faktorer som er med på å påvirke *fokuset på kvalitetsvurdering*. Man kan hevde at det nå skjer en utvikling der påvirkning fra internasjonale standarder for hva som er god utdanning gjør seg mer og mer gjeldende. Dermed kan man si at det internasjonale marked spiller en relativt stor rolle i utformingen av den nasjonale utdanningspolitikken. Datateknologi gjør det enklere for store nasjonale systemer å utøve kontroll. Det er mulig å utvikle billige overvåkningssystemer for monitorering av store organisasjoner. Dette bør sees på som kvantitative systemer som *kan* avdekke spesifikke problemstillinger, men som kun avdekker *deler* av komplekse systemer. Det er muligens belegg for å si at datateknologien kan være med på å overkjøre, og fjerne oppmerksomheten fra, det som ikke kan måles med kvantitative indikatorer.

Fokuseringen på indikatorer bør sees i sammenheng med både kvalitets- og evalueringsbølgen, NPM, internasjonalisering og teknologi, og kan derfor oppfattes som betydningsfull for spørsmålet om *hvorfor kvalitetsvurdering har utviklet seg til å bli et*

*sentralt element på den utdanningspolitiske dagsorden.* Man kan ane et økende krav til produksjon og effektivitet i kunnskapssamfunnet, noe som ser ut til å skape et behov for mer utdanning. Tradisjonelt sett kan man si at slike krav har blitt løst ved å gi mer ressurser til utdanningssystemet. Nå er det rimelig å hevde at virkeligheten er en annen: Ressursbruken ser ut til å reduseres, samtidig som det kan se ut som om det kreves mer utdanning for hver krone. Dette kan ha noe å si for styringen av utdanningssystemet, og man kan i lys av dette hevde at et behov for nytenkning vil vise seg. Hvis dette er tilfellet, kan man si at kontroll av kvalitet og ressursbruk blir fremtredende, og at konkurransen utvikler seg til å bli et av de viktigste virkemidlene. Det vil da bli fokusert på indikatorer som kan måle og sammenlikne, og incentivordninger av diverse art som kan belønne eller straffe innsatsen.

Det kan se ut som det i dag bare blir viktigere og viktigere for den utdanningspolitiske dagsorden å vise seg som moderne og fremtidsrettet, det vil si vise at de ”henger med på lasset” i utviklingen som skjer både nasjonalt og internasjonalt. Kvalitet og effektivitet kan dermed oppfattes som moteord på den utdanningspolitiske dagsorden. I et symbolperspektiv kan man hevde at NPM har utviklet seg til å bli en institusjonalisert oppskrift (Bendheim & Hansen 2001:270), den kan sies å være objektivert i det offentlige systemet og dermed tolkes som en filosofi som påvirker utviklingen på den utdanningspolitiske dagsorden.

*Kan dette sees på som en avideologisering av den utdanningspolitiske dagsorden?* Den høyreorienterte statsviteren Francis Fukuyama publiserte i 1989 en artikkel med tittelen *The End of History* (Telhaug & Mediås 2003:274). Her hevdet han at den ideologiske striden mellom venstresiden og høyresiden var i ferd med å avsluttes. Historien hadde kommet til veis ende etter sammenbruddet i de kommunistiske statene i øst. Fukuyama påstod at alle avanserte land hadde adoptert det liberale demokratiet som politiske institusjon, og de hadde godtatt en markedsorientert økonomi med integrering i en global og kapitalistisk arbeidsdeling (ibid). Man kan kanskje hevde at Fukuyamas tolkning også kan benyttes for å forklare utviklingen på den utdanningspolitiske dagsorden og fokuset på kvalitetsvurdering og NPM-filosofien. Dette fokuset kan sies å være det riktige å velge for å skape en identitet som kan oppfattes som rasjonell og som kan oppnå legitimitet. Behovet for en ny identitet på den utdanningspolitiske dagsorden kan sies å både komme fra en ekstern samfunnsmessig påvirkning og fra en intern

utvikling. Man kan i lys av dette si at skolen ikke lenger er et isolert segment som skiller seg ut fra den helhetlige politiske aktivitet. Det kan med andre ord hevdes at det har skjedd en politisering av segmentet utdanning som skaper et behov for nye orienteringer på den politiske dagsorden. Man kan oppfatte dette som at segmentet utdanning nå ligger på et mer åpent plan som skaper rom for nye aktører og eksterne drivere. Her kan man hevde at det er denne utviklingen som har skapt et rom for NPM-filosofiens elementer som etter hvert har utviklet seg til å bli oppskrifter for offentlig virksomhet, også for segmentet utdanning. Ut fra denne oppgavens fremstilling av NPM kan man oppfatte at det i denne filosofien, og dens fokus på blant annet kvalitet og evaluering, ligger en iboende retorikk som matcher en ny og passende identitet for de utfordringene som den utdanningspolitiske dagsorden møter i dagens samfunn.

### 6.1. NPM – en ”protected discourse”?

Med referanse til Bendheim og Hansen kan man hevde at NPM har utviklet seg til å bli en institusjonalisert oppskrift (Bendheim & Hansen 2001). Man kan muligens kritisere oppfatningen om at NPM er én oppskrift, man bør nok heller se på NPM som et sett av oppskrifter som tilhører samme ”filosofi”. I lys av en slik forståelse er det mulig å si at NPM-diskursen følger en entydighet i oppskrift og handling. Kan man hevde at denne diskursens argumenter er de eneste som legges til grunn *for* kvalitetsvurdering på den utdanningspolitiske dagsorden?

Det er problematisk å svare ”ja” på dette spørsmålet. Det er fortsatt flere diskurser, med helt andre utgangspunkt og orienteringer, som kan sies å være gjeldende. Spesielt kan man her se på en av de sentrale diskursene i *kvalitetsvurdering*, nemlig evalueringsdiskursen. Denne diskursen har utviklet seg gjennom forskningsaktivitet og teoridannelse, det er mulig å iakttå en *historie* i evalueringsdiskursen.

Evalueringsdiskursen kan hevdes å være en motsats til NPM-diskursen da førstnevnte representerer noe kritisk og reflektert, mens sistnevnte heller kan sies å representere noe entydig og oppskriftsmessig. Dermed er det rimelig å anta at argumentasjonen for eller mot kvalitetsvurdering ikke er låst, da NPMs oppskrifter antakeligvis vil bli utsatt for kritisk refleksjon fra en mer vitenskapelig tilnærming som trekker inn historiske og



verdimessige elementer inn i debatten. Her kan det være legitimt å vise til et eksempel som omhandler tolkning og anvendelse av evalueringsdata, noe som videre kan forstås som en demokratidebatt. For å belyse denne debatten kan man henvise til resultater fra for eksempel PISA-undersøkelsen og de oppgavene en evaluator bør ha/ikke bør ha i arbeidet etter at dataene er samlet inn.

Resultatene fra PISA kan brukes som del av kontroll og del av informasjon for utvikling, men man kan hevde at resultatene sier for lite om hvorfor Norge plasseres ”midt på treet”. Resultatene fra undersøkelsen kan brukes på den utdanningspolitiske dagsorden for å sette i gang tiltak som man *tror* fungerer. Bemerkelsesverdig er det at det ikke finnes noen forskning på disse resultatene ennå (om *hvorfor* midt på treet). Dermed kan man forstå at resultatene kan bli brukt i et maktpolitisk spill der politikere har mulighet til å plukke fritt fra resultatene og tolke dem i en retning som er med på å støtte *det gjeldende politiske programmet*. Det vil si at resultatene anvendes i en politisk prosess som kanskje ikke er ønskelig på sikt, men som kan oppnå legitimitet her og nå.

Dette betegner en demokratisk debatt som etter min mening er sentral i denne sammenheng, og ikke minst viktig for den utdanningspolitiske dagsorden. Sentrale spørsmål kan her være: Hvem eier resultatene? Hvem har rett til å tolke dataene? Skal dataene offentliggjøres? ... uten videre tolkning?

Den demokratiske debatten kan settes i gang ved å ta utgangspunkt i Haugs syv dimensjoner for evaluering (Haug 2003b). De syv dimensjonene kan oppfattes som en drøfting omkring hvor langt evaluator skal gå i den videre prosessen etter at dataene er samlet inn. I bunn og grunn handler det om hvilket ansvar evaluator står overfor i en evalueringssituasjon.

Den første dimensjonen handler om å *beskrive*. Vi kan her vise til et eksempel fra evalueringen av Reform 97 hvor det kommer frem at jenter får gjennom gående bedre karakter enn guttene i norsk (Haug 2003a). Dette er en beskrivelse. Tidligere har dette blitt sett på som evaluering, dvs. at det er nok å legge frem dataene. Haug poengterer at dette bare er et utgangspunkt, det er kun en beskrivelse, ikke en evaluering.

Beskrivelsene utdateres raskt, alene har de ikke lang gyldighet. Det må mer til for å kalle dette en evaluering.

*Forklaringsdimensjonen* bringer oss et stykke videre. Her står metodikken i sentrum; det handler om å forklare årsaker, sammenheng og forstå hva beskrivelsene egentlig handler om. Man må forklare hvilke metoder som er brukt for å si noe om hvorfor man har kommet frem til akkurat dette resultatet, og ikke noe annet. I denne dimensjonen handler det om å reflektere over hvilke spørsmål som har blitt stilt, det handler om ærlighet, det vil si vise hvilke interesser som ligger bak evalueringen. I motsetning til beskrivelsene, er forklaringene/analysene holdbare, de utdateres ikke på samme måte som beskrivelsene/dataene. Et eksempel på dette kan fremmes ved å henvise til Freud: diskusjonsmomentene er fortsatt gyldige, men det er ikke dataene.

Forklaringsdimensjonen møter ofte motstand innenfor evalueringsfeltet. Det finnes sterke aktører i feltet som mener at forklaringen ikke har noe å si. Oppdragsgiver er først og fremst opptatt av om det fungerer bra eller ikke: "Vi velger dette fordi det har evalueringsdataene vist at fungerer".

Den tredje dimensjonen er *vurdering*. De fleste oppdragsgivere forventer "tommelen opp/ned". Noen ganger er dette mulig (for eksempel i medisinforskning), andre ganger må man gi svar i analyser, og videre analyserer. Det er et problem innenfor evaluering at verdivurderingen ofte mangler, mange lider av det Scriven kaller *value fobia*<sup>58</sup>. Her viser det seg hvor viktig det er med en velfundert programteori og eksplisitte kriterier for ønsket resultat. Vi kan også si at det her ligger et skille mellom evaluering og forskning. I forskning er det ikke forventet at man skal verdsette, men det er det i en evaluering.

*Informasjon* er den fjerde dimensjonen. Denne dimensjonen henger nøye sammen med spørsmålet om hvem som eier resultatene og dermed om hvordan resultatene blir brukt. Politikerne, som ofte er oppdragsgivere, plukker ofte ut løse deler av resultatet og bruker disse retorisk. Vi kan her vise til AEAs<sup>59</sup> etiske regelverk der det heter seg at *alle* evalueringsresultater skal offentliggjøres. Dette er et prinsipp i et demokratisk perspektiv; alle skal ha tilgang på resultatene. Publiseringen innebærer alltid et dilemma: Er det forskerne som eier resultatene, og kan publisere dem som de vil? Eller

---

<sup>58</sup> "Value fobia" defineres som "a fobia against estimating value" (Scriven 1991a).

<sup>59</sup> American Evaluation Association

eies resultatene av oppdragsgiver (for eksempel et departement), og det er deres egen interesse som bestemmer hva som skal publiseres, og hvordan?

Den femte evalueringsdimensjonen er *forhandling*. Når man publiserer resultatene oppstår det ofte en offentlig diskusjon. I hvor stor grad skal en evaluator gå inn i denne diskusjonen? Man kan her argumentere både for og mot: Evaluator burde ikke gå inn i den offentlige diskusjonen da det er opp til hvert enkelt medlem/gruppe i samfunnet å tolke resultatene og uttale seg på bakgrunn av denne oppfatningen. Dette er en offentlig debatt som er til for alle, ikke kun for ekspertene. Det er ikke hensiktsmessig i et demokratisk samfunn å skape enighet. På den andre siden kan man hevde at evaluator burde delta i debatten for å klargjøre eller oppklare eventuelle feil og feiltolkninger. Problemet med enhver innblanding er at evaluator da blir en politisk aktør, noe som per definisjon ikke er ønskelig.

*Forslag*, det å foreslå, er den sjette dimensjonen. Spørsmålet som stilles her er: Skal evaluator foreslå tiltak, det vil si vise hva bestiller bør gjøre? Forslag til tiltak er alltid problematisk fordi evaluator aldri har det fulle og hele bilde av organisasjonen (ser ikke det politikerne eller byråkratene gjør). De som har analysert forslag til tiltak ser at de fleste forslagene kunne vært fremstilt uavhengig av evalueringen. Det viser seg at evaluators *for-dommer* til institusjonen/organisasjonen spiller en større rolle enn de faktiske evalueringsresultatene gjør. Det finnes ikke mange slike metaevalueringer, men tendensene er at forslag ofte ikke samsvarer med evalueringsresultatene.

Den siste dimensjonen er *handling* og dreier seg om i hvor stor grad evaluator skal bli en del av tiltakene/sloss på barrikadene. Det handler kort og godt om det er riktig at evaluator skal bli en politisk aktør.

Det kan være fristende å kritisere en del av evalueringsteoretikere for å ikke delta i debatten etter at evalueringsresultatene er lagt frem. Dette kan, etter min mening, sees på som en form for verdifobi; forskerne unngår, på det faglige grunnlag de har, å ta standpunkt til det feltet de undersøker. Dette betyr at evalueringsforskerne er med på å godta og legitimere systemet som det *er*. Hvis de tar et verdistandpunkt kan de komme i en situasjon der de kritiserer det etablerte systemet, det vil si fremsetter en samfunnskritikk som ikke er ønskelig innenfor det (makt-)systemet de opererer i. Dette

systemet viser til makthaverne, de styrende på den utdanningspolitiske dagsordenen, det vil si oppdragsgiverne, og den politiske formuleringsarenaen. Det handler dermed om hvilke diskurser som vinner frem i det diskursive felt, hvem som fremhever hvilke diskurser og hvilke diskurser som oppnår hegemoni.

”PISA – et skjevt tårn?” er overskriften i kapittel 14.2.2 i den norske PISA-rapporten *Godt rustet for framtida? Norske 15-åringers kompetanse i lesing og realfag i et internasjonalt perspektiv* (Lie m.fl. 2001). Her reflekterer de ansvarlige forskerne fritt over PISA-undersøkelsen og de norske resultatene. Det er verdt å merke seg deres refleksjoner over hvilke kompetanser som er viktige i morgendagens samfunn og hvordan PISA/OECD velger ut sitt rammeverk. Svein Lie m.fl. forteller oss at rammeverket for PISA tar utgangspunkt i hva som anses som viktige kompetanser i morgendagens samfunn. De kommenterer dette med å si: ”Det sikreste vi kan si om morgendagens samfunn, er at det blir annerledes.” (op.cit:280). Deretter understrekes det at OECD er en organisasjon av rike land, med økonomisk vekst som mål. Dette er en svært viktig refleksjon å ta med seg når man arbeider med og tolker resultater fra en undersøkelse som PISA. Prinsippene som legges til grunn for rammeverket i PISA vil nødvendigvis være preget av et økonomisk siktemål. Det er utdanning for arbeidslivet som står i fokus, mens perspektiver for fritid og kulturliv ikke blir behandlet/fokusert på. Dermed ser vi at rammeverket er preget av et nokså instrumentelt syn på utdanning: Kunnskap er viktig fordi den skal brukes i arbeidslivet, i samfunnets tjeneste. I PISAs rammeverk er altså kunnskap for kunnskapens egen skyld, kunnskap i et dannelsesperspektiv, ikke fremtredende (ibid).

Refleksjoner omkring dette bildet bør være med når en går løs på å tolke resultatene fra en undersøkelse som PISA – og før man setter i gang med spesifikke tiltak. Det ligger muligens et kritisk punkt i å la resultatene stå for seg selv uten videre forklaringer og velfunderte tolkninger. Det kan resultere i at resultatene blir brukt på den utdanningspolitiske dagsorden på en sådan måte at det ender i tiltak som bygger på et skjult grunnlag. Da er det kanskje mulig å se en utvikling mot det Skjervheim kaller *det instrumentalistiske mistaket*<sup>60</sup>.

---

<sup>60</sup> Mer om det instrumentalistiske mistaket kan finnes i Skjervheim, H. 1996. Deltakar og tilskodar og andre essays. Oslo: Aschehoug (Det instrumentalistiske mistaket, side 241-250).

Den økonomiske instrumentalismen og nyttetenkningen som man kan oppfatte som fremtredende i NPM, kan diskuteres ut fra Skjervheims bruk av det aristoteliske skillet mellom *poiesis* og *praxis*, eller mellom pragmatiske (tekniske) og praktiske handlinger. Telhaug og Mediås skriver at pedagogikken i 1990-årene ble forstått som *poiesis* eller pragmatisme av politikerne (Telhaug & Mediås 2003). Dette medførte at kunnskapen ble vurdert ut fra mål som befant seg utenfor kunnskapen selv. Det er dette som skjer når kunnskapen vurderes strategisk, ut fra de resultater som kan fremme økonomiske vekst (ibid). Denne tankegangen kritiseres ut fra et syn der kunnskapen forstås som praksis, i betydningen at handlingene skjer ut fra grunnleggende prinsipper som alle andre også skal kunne handle ut fra.

I lys av det som her har blitt omtalt som en demokratisk debatt, kan man hevde at det ikke er legitimt å snakke om at NPM har utviklet seg til å bli en *protected discourse* eller at det kun er NPMs argumenter som gjør seg gjeldende i argumentasjonen *for* kvalitetsvurdering på den utdanningspolitiske dagsorden. Man kan fortsatt hevde at det er diskursive kamper som gjør seg gjeldende på den utdanningspolitiske dagsorden, som videre gjør det problematisk å fastlåse en bestemt meningskonstruksjon. Det kan se ut som om den demokratiske debatten om blant annet hvem som eier evalueringsdata og hvem som skal tolke disse, fortsatt vil prege den utdanningspolitiske dagsorden i tiden fremover. Dermed er det rimelig å anta at kampen i skjæringspunktet mellom den teoretiske, vitenskapelige arena og den samfunnsmessige, politiske arena vil vedvare.

I det neste, og siste, kapittelet vil jeg stille spørsmål om det er mulig å snakke om en institusjonalisering av kvalitetsvurderingssystemet og om dette kan tolkes som en bevegelse mot en markedsstyrt utdanningsideologi.

## 7. INSTITUSJONALISERING

”In the tradition of modern sociology, from Weber to Parsons and Luhmann, it makes sense to describe the process of *institutionalization* as the main force shaping the structures of modern Western societies. The theories may differ somewhat in what they see as the driving forces behind this process (functional or other types of differentiation, belief systems, etc.), but they agree on the fact that we organize most of our social life with help of institutionalizations which give our activities more or less stable frames of reference.” (Hopmann 2003:115).

Kan man hevde at den utdanningspolitiske dagsorden i dag er noe annet enn tidligere? Er fokuset på kvalitetsvurdering med på å skape en ny identitet? Er argumentasjonen i dag en annen enn tidligere, og oppnår den utdanningspolitiske dagsorden rasjonalitet og legitimitet til denne argumentasjonen *for* kvalitetsvurdering? Er det legitimt å hevde at man nå er på vei mot en markedsstyrt utdanningsideologi?

Disse spørsmålene handler i bunn og grunn om institusjonalisering og hvilken betydning et nytt kvalitetsvurderingssystem vil få i utdanningen i fremtiden. Man kan muligens hevde at kvalitetsvurdering, og spesielt fokuset på evalueringens funksjon som problemløser, bringer nye styringsverktøy på banen<sup>61</sup>. Det er rimelig å anta at disse har betydning for politikkkutforming og kjemper for å vinne legitimitet blant fagfolk og i offentligheten. Formålet med å oppnå rasjonalitet og legitimitet kan sies å være at kvalitetsvurderingssystemet blir institusjonalisert, eller objektivert hvis man skal referere til diskursteoriens termer. Med henvisning til diskursteorien vil dette si at man ser på kvalitetsvurdering som et nødvendig verktøy i skolesystemet, man tar det for gitt og ”glemmer” at omverdenen alltid konstituerer seg i makt og politikk. Dermed kan det tyde på at det å skape et rasjonelt verktøybilde av symboler gjør seg gjeldende i organisasjonssamfunnet for å nå dette målet. Det handler da om hvilke diskurser som vinner frem, hvilke argumenter som fremsettes for å gi kvalitetsvurderingen rasjonalitet og oppnå legitimitet.

Med henvisning til kapittel 3.3 kan institusjonalisering tolkes som et begrep som brukes for å vise hvilke prosesser som fører til at noe i en organisasjon gradvis fremstår som et regelliknende faktum, noe som tas for gitt som en effektiv og riktig måte å organisere på. Vi kan også vise til diskursteoriens objektivitetsbegrep som grunnlag for hvordan

---

<sup>61</sup> For eksempel resultatstyring og styring gjennom incentiver og konkurranse.

noe institusjonaliseres i en organisasjon<sup>62</sup>. Det er ikke et mål her å beskrive hvordan kvalitetsvurdering har blitt institusjonalisert på den utdanningspolitiske dagsorden og i skolesystemet som helhet – det har jeg ikke grunnlag for å si noe om<sup>63</sup>. Men jeg kan vise til hvilke tendenser man kan se på den utdanningspolitiske dagsorden og hvilken argumentasjon som legges til grunn *for* kvalitetsvurdering på den utdanningspolitiske dagsorden.

### 7.1. Hvilke diskurser vinner hegemoni?

Man kan oppfatte NPM-filosofien som en advokat *for* kvalitetsvurdering, med sin tro på næringslivets verdier og sitt rasjonelle bilde av politikkutforming. Hvis det er slik at NPM-filosofiens argumenter har vunnet hegemoni på den utdanningspolitiske dagsorden, skulle man tro at disse argumentene ville komme klart frem i nyere, sentrale politiske dokumenter. Med andre ord kan man da iakttå blant annet relativt hyppig bruk av økonomiske begreper og referanser til oppskrifter som har vist seg å fungere på det private marked og som der har fremmet økonomisk vekst. Videre i dette bildet kan man hevde at konkurransen vil vise seg som virkemiddel, med incentiver som skal øke kvaliteten i hele skolesystemet. I tillegg vil det vise seg et behov for desentralisering av mange av skolens oppgaver, samtidig som det sentrale nivået har en kontrollerende hånd over det hele. Man kan da hevde at en argumentasjon *for* kvalitetsvurdering bygger på økonomiske begreper og (inter-)nasjonale kvalitetsstandarder. Til sammen skal dette skal skape et bilde av kvalitetsvurdering som et nyttig og nødvendig verktøysystem i skolen. Hvis dette bildet er å finne i nye, sentrale dokumenter, kan man tolke det som en utvikling mot resultatstyring der man kan si at konkurransen er fremtredende.

Dette hegemoniet inkluderer også en sterk tro på evaluering som løsning på skolens mange utfordringer. Det å vise til evalueringsresultater kan være med på å gi kvalitetsvurdering både rasjonalitet og legitimitet. I et symbolperspektiv kan man hevde at evalueringsresultater som blir fremsatt som klare bilder på hvordan skolesystemet

---

<sup>62</sup> Se kapittel 2.7.1.

<sup>63</sup> Etter min mening er det for tidlig å snakke om at kvalitetsvurdering er institusjonalisert som et system i skolen.

fungerer, er symboler som blir omformet til rasjonelle verktøy på den utdanningspolitiske dagsorden.

## 7.2. Den blå bølge inn på den utdanningspolitiske dagsorden?

Kan man hevde at fokuset på kvalitetsvurdering er med på å åpne for den tidligere nevnte blå bølgen på den utdanningspolitiske dagsorden?

Man bør muligens stille dette spørsmålet på en annen måte. Man kan kanskje heller hevde at det er den blå bølge og NPM-filosofiens inntog på offentlig sektor som er med på å skape et behov for nye (styrings-)verktøy som kvalitetsvurdering. For å ta stilling til dette spørsmålet, må man først vurdere om det er grunnlag for å hevde at den blå bølge, NPM-filosofien, har vunnet hegemoni på den utdanningspolitiske dagsorden.

Oppgaven har fokusert på NPMs argumenter for en mer effektiv offentlig styring og hvilken påvirkning denne filosofien har og kan få for den utdanningspolitiske dagsorden. Om denne filosofien kan sies å være fremtredende når det nå argumenteres for en ny læreplan, vil være gjenstand for diskusjonen avslutningsvis i denne oppgaven.

St.meld. nr. 30 (2003-2004) *Kultur for læring* ble offentliggjort 2. april 2004 og er Utdannings- og forskningsdepartementets (under regjeringen Bondevik II) forslag til utviklingen av en ny læreplan og ”ny” styring av grunnopplæringen<sup>64</sup>. Kan man med rett hevde at kvalitetsvurdering er et ”barn av NPM-filosofien” og at fokuset på kvalitetsvurdering dermed bringer den blå bølge inn på den utdanningspolitiske dagsorden? Ved å vise til *Kultur for læring* og hvilke argumenter som der brukes, kan man kanskje belyse dette spørsmålet. Man kan da lete etter hvilken argumentasjon som legges til grunn for å skape rasjonalitet og oppnå legitimitet til kvalitetsvurdering på den utdanningspolitiske dagsorden. Kan man øyne at det er en institusjonalisering av kvalitetsvurdering og NPMs elementer på vei?

---

<sup>64</sup> Ordet grunnopplæring betyr grunnskole og videregående opplæring (Stmeld. nr. 30 2003-2004).



Det kan se ut som om fokuset på kvalitetsvurdering, og utviklingen av et system for kvalitetsvurdering, er fremtredende i den nye stortingsmeldingen. I innledningskapittelet henviser meldingen til kvalitetsutvalgets arbeid og deres delinnstilling NOU 2002:10 *Førsteklasses fra første klasse. Forslag til rammeverk for et nasjonalt kvalitetsvurderingssystem av norsk grunnopplæring*. Meldingen forteller videre at Stortinget ga sin tilslutning til utviklingen av et nasjonalt kvalitetsvurderingssystem. Dette kvalitetsvurderingssystemet skal blant annet inneholde nasjonale prøver i sentrale fag og nettstedet ”skoleporten.no”.

Kapittel 3 i meldingen viser til at det nasjonale kvalitetsvurderingssystemet vil gi skolen kunnskap som kan brukes som utgangspunkt for endring og utvikling. Det nasjonale systemet som kvalitetsvurdering og -utvikling beskrives som en viktig del av et større systemskifte. Regjeringen mener det er behov for en styring som i større grad enn før baserer seg på nasjonale mål, tydelig ansvars plassering og økt lokal handlefrihet<sup>65</sup>. Systemskiftet beskrives av regjeringen med ordene *frihet, tillit og ansvar*, og det baseres på grunnprinsipper om ”klare nasjonale mål, kunnskap om resultater i vid forstand, tydelig ansvars plassering, stor lokal handlefrihet og et godt støtte- og veiledningsapparat”. Videre i kapittel tre heter det:

”Det nasjonale kvalitetsvurderingssystemet vil gi informasjon om læringsresultater, ressurser, læringsmiljø og trivsel. Dette er først og fremst et virkemiddel for den enkelte skole og skoleeier til å skaffe kunnskap som grunnlag for utvikling. Nettstedet skoleporten.no skal bidra til åpenhet og legge bedre til rette for at skolene selv, skoleeiere, foresatte, elever og andre interesserte kan engasjere seg i utviklingen av skolen.” (St.meld.nr.30 2003-2004:26).

Det nasjonale kvalitetsvurderingssystemet begrunnes med at både staten, skoleeier og skolen selv har behov for å ha god oversikt over det som skjer i skolen. Det blir stilt krav fra skoleeier og statlige myndigheter om ulike rapporteringer som i følge stortingsmeldingen har tre formål. For det første begrunnes rapporteringen i behovet for regelverksoppfølging og gir et grunnlag for tilsyn. Det andre formålet viser til at rapporteringen gir et grunnlag for utviklingsarbeidet gjennom å avdekke områder for forbedring. Til slutt er rapporteringskrav begrunnet i behovet for ettersyn og kontroll med økonomiske støtteordninger (ibid).

---

<sup>65</sup> Jamfør accountability-begrepet.

I kapittel fire i meldingen blir ”broen” i nåværende læreplan (L97/S) utsatt for kritikk. Departementet mener mye av innholdet er vurderinger og bestemmelser som angår planlegging og tilrettelegging av opplæringen. ”Broen” beskriver også hvordan opplæringen best kan gjennomføres. Denne formuleringen er ikke departementet enig i. I meldingen argumenteres det for at dette er et lokalt, profesjonelt ansvar, og at det derfor i all hovedsak bør overlates til skoleeier/skoler å velge hvordan opplæringen planlegges, tilrettelegges og gjennomføres. Som en *erstatning* til ”broen” definerer departementet noen grunnleggende prinsipper og krav som skal prege og forplikte alle skoler og øvrige opplæringssteder. Disse defineres som ”Rammeverk for kvalitet – Skoleplakaten” og blir beskrevet i 10 krav som skolen og lærebedriften forplikter seg til. Departementet foreslår at Skoleplakaten fastsettes som forskrift slik at prinsippene kan bli gjenstand for statlig tilsyn (op.cit:37). Prinsippene skal sees i sammenheng med det nasjonale kvalitetsvurderingssystemet (ibid).

Tidligere i denne oppgaven har det blitt hevdet at kvalitetsvurdering kan sees på som et desentraliseringstiltak. Og det har også blitt argumentert for at NPM-filosofien vektlegger en form for desentralisering av oppgaver, samtidig som det utvikles kvalitetskontrollmekanismer. Kan man finne liknende tendenser i *Kultur for læring*?

Meldingen fremsetter argumenter til fordel for desentralisering av planlegging, tilrettelegging og gjennomføring av opplæringen<sup>66</sup>. Det argumenteres for at skoleeier, skoleledere, lærere, elever og lærlinger bør få stor frihet til å organisere seg innenfor nasjonale mål (op.cit:8). Overfor ble det referert til det nye styringssystemets grunnprinsipper hvor *stor lokal handlefrihet* var et av prinsippene. Dette argumenteres med et skifte fra forestillingen om at staten kan skape et likeverdig skoletilbud gjennom detaljregulering og -styring, til en mer desentralisert styringsfilosofi der man har tillit til at den enkelte lærer, skoleleder og skoleeier selv har de beste forutsetningene for å vite hvordan god læring kan skapes og gjennomføres (op.cit:25). Med denne økningen i handlefrihet argumenteres det for at skoleeier gis både nye muligheter og større ansvar. Dette ansvaret henger ikke i løse luften. Staten skal føre tilsyn med at skoleeierne ivaretar sitt ansvar i henhold til opplæringsloven, og at elevenes rettigheter blir ivaretatt. ”Dette innebærer et behov for kunnskap og åpenhet om skolens evne til å nå sine mål”

---

<sup>66</sup> Jamfør accountability-begrepet.

(op.cit:26). Her kommer det nasjonale kvalitetsvurderingssystemet inn som en kontrollmekanisme fra statlig hold. Dermed kan man kanskje hevde at det nasjonale kvalitetsvurderingssystemet, slik det viser seg i *Kultur for læring*, er et desentraliseringstiltak med relativt sterke sentraliserte trekk.

Men kan man hevde at det er den blå bølge som bærer dette fokuset?

Meldingen kan oppfattes som uttrykk for en både-og filosofi der desentralisering åpner for den lokale kunnskapen, mens det man kan kalle sentraliseringstendensene og fokuset på få og klare mål som formuleres i standardiserte kvalitetskriterier, er noe som kan sees på som en bevegelse i retning av instrumentalisme. Dette kan igjen tolkes dit hen at man nå er på vei mot en klarere resultatstyring av skolen enn tidligere (på de områdene den utdanningspolitiske dagsorden mener er betydningsfull for den nasjonale økonomien), noe som gir et bilde av et produktfokus. ”Produktene”, det vil si resultatene, skal dokumenteres og ”plottes” inn som del av kvalitetsvurderingssystemet og skal virke som en utviklings- og forbedringsdriver for den enkelte skole. En del av dette systemet har som sagt fått navnet ”skoleporten.no” og betegnes som en av skole-Norges nyvinninger<sup>67</sup>. Dette nettstedet skal inneholde opplysninger om alle skoler i Norge og skal fungere som et verktøy for kvalitetsvurdering og kvalitetsutvikling. Skoleporten.no inneholder blant annet kartleggingsresultater og veiledningsinformasjon for de fire kvalitetsområdene ressurser, læringsmiljø, læringsutbytte og gjennomstrømming<sup>68</sup>.

Det er kanskje mulig å tolke denne ”nyvinningen” som et kommunikasjons- og informasjonsverktøy som etter hvert kan utvikle seg til å bli en egen styringsfaktor for skolene<sup>69</sup>. Dette systemet med offentliggjøring av nasjonale prøver har i den offentlige debatten blitt kritisert for å likne på det man i dag kaller ”test and tell”. I en avisartikkel skriver journalist Lars Kluge om *skolens fallgruber* (Kluge 2004). Han henviser der til at det amerikanske skolesystemet og deres ”test and tell”-ideologi. Denne kan oppfattes som en ideologi og et styringsverktøy som går ut på å teste *nøyaktig* hva elevene kan (ibid). Det er en selvfølge at resultatene offentliggjøres i det amerikanske systemet.

---

<sup>67</sup> Se [www.skoleporten.no](http://www.skoleporten.no)

<sup>68</sup> Et fjerde kvalitetsområde skisseres her, i tillegg til de tidligere nevnte fra Kvalitetsutvalgets innstilling (se kapittel 3.1.1).

<sup>69</sup> Jamfør kapittel 4.3, side 58-59.

Kluge er opptatt av *fallgrubene* man kan gå i ved å følge en slik ideologi i Norge. Han skriver: ”Faren er at disse målene blir skolens egen målestokk” (ibid). Og han mener at drillingen av elevene for å oppnå best mulige resultater på de standardiserte testene, kan komme i veien for selve læringen. I lys av dette kan man her referere til den tidligere nevnte strategien ”teaching to the test”. Faren ved en slik offentliggjøring kan i tillegg være at de gode elevene og lærerne flykter fra de skolene som kommer dårlig ut på de nasjonale prøvene, altså de skolene som ender på bunnen i den nasjonale sammenlikningen. Ved en utvikling som dette kan man hevde at kommunikasjon og informasjon har skapt en gjennomsiktighet (transparency) i skolesystemet som er med på å styre skolen.

Med henvisning til *Kultur for læring* kan det være problematisk å hevde at det er næringslivets verdier og den blå bølge som er bærer av fokuset på kvalitetsvurdering. I meldingen argumenteres det hele tiden til dannelsesperspektivet og individet når det henvises til det man kan kalle mer instrumentalistiske begrunnelser. Dermed bør man kanskje heller tolke meldingen som en fortsettende diskursiv kamp der det ennå ikke er noe hegemoni å snakke om. Med en slik forståelse kan man si at diskursen ikke er lukket eller avsluttet. Meldingen gir muligheter for hegemonikamp da formuleringene kan oppfattes som nok uklare og forsiktige. Overskriften *Kultur for læring* kan muligens tolkes som et retorisk grep for å skape et fokus på begrepene læring og kultur i stedet for på begreper som konkurranse og kontroll.

### 7.3. Et kompleks fenomen

Kvalitetsvurdering bør sees på som et kompleks fenomen som ikke med én enkel forklaring kan skape en forståelse for hva dette innebærer som evaluering eller hvilke konsekvenser det kan ha for utdanningssystemet. Målet med denne oppgaven har vært å tydeliggjøre denne kompleksiteten og utvide forståelsen av den – og av fenomenet kvalitetsvurdering. Man kan hevde at det den dag i dag ikke er mulig å fremstille kvalitetsvurdering som et enkelt fenomen som uproblematisk løser skolens nåtidige og fremtidige utfordringer. Både den historiske fremstillingen og samfunnsutviklingen kan belyse dette.

Utfordringen videre bør være å arbeide for å skape et klarere bilde av hva man vil med kvalitetsvurderingsystemet, og å hele tiden vurdere de langsiktige konsekvensene et slikt system kan ha. Suksessen til de nye verktøyene som tas i bruk kan sies å være avhengig av at verktøyene ikke blir ”døgnfluer” som hele tiden skiftes ut og aldri implementeres nedover i systemet. De må oppnå legitimitet i hele systemet, de må vise seg som verktøy som kan ”make a difference”.

Denne oppgaven har forsøkt å belyse fenomenet kvalitetsvurdering og dets inntog på den utdanningspolitiske dagsorden, og forsøkt å skape en forståelse for fenomenets kompleksitet. Hvordan et kvalitetsvurderingssystem vil virke, og hvilken innvirkning det vil ha på opplæringen og hverdagen på den enkelte skole, er vanskelig å si noe om. Man bør hele tiden være oppmerksom på hvilke feller et slikt system kan ha, og dermed hvilke konsekvenser det kan få i fremtiden.

En siste bemerkning: Når man nå kan si at fenomenet kvalitetsvurdering på den utdanningspolitiske dagsorden har utviklet seg til et system med både form og innhold<sup>70</sup>, er det etter min mening ikke slik at man bør oppfatte det som et ”ferdig” kvalitetsvurderingssystem som vil fremme kvalitet og positiv utvikling i fremtidens skole. Man vet ikke hva som vil være kvalitet og positiv utvikling i fremtidens skole, og det kan tenkes at dette systemet ikke overlever *morgendagen*. Man bør hele tiden være oppmerksom på, og reflektere over, hvilke konsekvenser dette systemet har, og kan få. Det er mulig at et slikt system fører med seg utilsiktede, og kanskje langvarige, konsekvenser. Disse kan selvfølgelig virke i både positiv og negativ retning, men poenget bør være at man kontinuerlig vurderer om dette gjør skolen til et bedre sted å være og et bedre sted å lære. Med referanse til Dahler-Larsen og Krogstrup kan man si at evalueringer er med på å endre og skape den virkeligheten vi lever i: ”Evalueringer beskriver ikke bare virkeligheten. De former den.” (Dahler-Larsen & Krogstrup 2001a:243).

---

<sup>70</sup> Se St.meld. nr. 30 (2003-2004) og [www.skoleporten.no](http://www.skoleporten.no)

## LITTERATUR

Albæk, E., Dahler-Larsen, P., Foss Hansen, H., Krogstrup, H.K., Larsen, F., Lassen, M. & Winter, S. 2001. Evalueringernes fremtid – fremtidens evalueringer. I: Dahler-Larsen, P. & Krogstrup, H.K. (Red.) 2001. *Tendenser i evaluering*. Odense: Syddansk universitetsforlag.

Bauman, Z. 1995. *Life in Fragments: Essays in Postmodern Morality*. Cambridge: Polity Press.

Beck, U. 1994. The Reinvention of Politics: Towards a Theory of Reflexive Modernization. I: Beck, U., Giddens, A. & Lash, S. 1994. *Reflexive Modernization. Politics, Tradition and Asthetics in Modedern Social Order*. Cambridge: Polity Press.

Bendheim, G. & Hansen, B.T. 2001. Reformer i norsk høgre utdanning – på vei mot New Public Management? I: Busch, T., Johnsen, E., Klausen, K.K., Vanebo, J.O. (Red.) 2001. *Modernisering av offentlig sektor. New Public Management i praksis*. Oslo: Universitetsforlaget.

Berger, P.L. & Luckmann, T. 1992. *Den samfundsskabte virkelighed. En videnssociologisk afhandling*. 2. ed. Viborg: Nørhaven A/S.

Bernstein, D.J. 1999. Comments on Perrin's "Effective Use and Misuse of Performance Measurement". *American Journal of Evaluation*, Vol. 20, No. 1, 1999, 85-93.

Busch, T. 2001. Konkurransen, økonomi og effektivitet. I: Busch, T., Johnsen, E., Klausen, K.K., Vanebo, J.O. (Red.) 2001. *Modernisering av offentlig sektor. New Public Management i praksis*. Oslo: Universitetsforlaget.

Carroll, L. 2000. *Through the Looking-glass*. Oxford: Oxford University Press.

Dahler-Larsen, P. 2001a. *Den rituelle refleksjon – om evaluering i organisasjoner*. Odense: Odense Universitetsforlag.

Dahler-Larsen, P. 2001b. Organisationalisering af evaluering. I: Dahler-Larsen, P. & Krogstrup, H.K. (Red.) 2001. *Tendenser i evaluering*. Odense: Syddansk universitetsforlag.

Dahler-Larsen, P. & Krogstrup, H.K. 2000. Evaluering og institutionelle standarder: Nyinstitutionelle betragtninger af evaluering som vor tids ritual. *Politica*, Vol. 32, No. 3, 283-299.

Dahler-Larsen, P. & Krogstrup, H.K. 2001a. Evalueringens konstitutive virkninger. I: Dahler-Larsen, P. & Krogstrup, H.K. (Red.) 2001. *Tendenser i evaluering*. Odense: Syddansk universitetsforlag.

Dahler-Larsen, P. & Krogstrup, H.K. 2001b. Tendenser i evaluering – en introduktion. I: Dahler-Larsen, P. & Krogstrup, H.K. (Red.) 2001. *Tendenser i evaluering*. Odense: Syddansk universitetsforlag.

Eco, U. 1990. *The Limits of Interpretation*. Bloomington, Indiana: Indiana University Press.

Foss Hansen, H. 2001. Rejsende praksis: Hvilke forhold påvirker udviklingen af evalueringspraksis på forskellige politikområder? I: Dahler-Larsen, P. & Krogstrup, H.K. (Red.) 2001. *Tendenser i evaluering*. Odense: Syddansk universitetsforlag.

Gall, J.P., Borg, W.R. & Gall, M.D. 1996. *Educational Research. An Introduction*. 6. ed. New York: Longman.

Giddens, A. 1997. *Modernitetens konsekvenser*. Oslo: Pax Forlag A/S.

Granheim, M., Lundgren, U.P. & Tiller, T. (Red.) 1990. *Utdanningskvalitet – styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole*. Oslo: Tano.

Granheim, M. 1990a. Forord. I: Granheim, M., Lundgren, U.P. & Tiller, T. (Red.) 1990. *Utdanningskvalitet – styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole*. Oslo: Tano.

Granheim, M. 1990b. Bakgrunnen for EMIL-prosjektet. I: Granheim, M., Lundgren, U.P. & Tiller, T. (Red.) 1990. *Utdanningskvalitet – styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole*. Oslo: Tano.

Haug, P. 1998. Linking evaluation and reform strategies. *New directions for evaluation: A publication of the American Evaluation Association*, no 77, 5-20.

Haug, P. 2003a. *Evaluering av Reform 97. Sluttrapport frå styret for Program for evalueringen av Reform 97*. Norges forskningsråd

Haug, P. 2003b. The Research Council of Norway evaluating Reform 97. I: Haug, P. & Schwandt, T.A. (Red.) 2003. *Evaluating educational reforms. Scandinavian perspectives*. Greenwich, Connecticut: Information Age Publishing.

Henmo, O. 2003. Kvalitet, kvalitet, kvalitet. I: *Aftenposten*. 16.08.03.

Holme, I.M. & Solvang, B.K. 1998. *Metodevalg og metodebruk*. 3. utg. Oslo: Tano.

Hopmann, S. 2003. On the Evaluation of Curriculum Reforms. I: Haug, P. & Schwandt, T.A. (Red.) 2003. *Evaluating educational reforms. Scandinavian perspectives*. Greenwich, Connecticut: Information Age Publishing.

House, E.R. 1990. Desentralisert evaluering i Norge. I: Granheim, M., Lundgren, U.P. & Tiller, T. (Red.) 1990. *Utdanningskvalitet – styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole*. Oslo: Tano.

House, E.R. 1993. *Professional evaluation: social impact and political consequences*. Newbury Park: SAGE Publications.



Judd C.M., Smith, E.R. & Kidder, L.H. 1991. *Research Methods in Social Relations*. 6. ed. Orlando: Harcourt Brace & Company.

Karlsen, G.E. 2002. *Utdanning, styring og marked: norsk politikk i et internasjonalt perspektiv*. Oslo: Universitetsforlaget.

Klausen, K.K. 2001. New Public Management – en fortolkningsramme for reformer. I: Busch, T., Johnsen, E., Klausen, K.K., Vanebo, J.O. (Red.) 2001. *Modernisering av offentlig sektor. New Public Management i praksis*. Oslo: Universitetsforlaget.

Kluge, L. 2004. Skolens fallgruber. I: *Aftenposten*, 15.04.04.

Kvale, S. 1990. Kunnskapsevaluering og maktdesentralisering. I: Granheim, M., Lundgren, U.P. & Tiller, T. (Red.) 1990. *Utdanningskvalitet – styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole*. Oslo: Tano.

Kvale, S. 1997. *Det kvalitative forskningsintervju*. Oslo: Ad Notam Gyldendal.

Lane, J.-E. 1995. *The Public Sector. Concepts, Models and Approaches*. 2. ed. London: Sage Publications.

Lie, S., Kjærnsli, M., Roe, A. & Turmo, A. 2001. *Godt rustet for framtida? Norske 15-åringers kompetanse i lesing og realfag i et internasjonalt perspektiv*. Oslo: Acta Didactica.

Lindensjö, B. & Lundgren, U.P. 1986. *Politisk styrning och utbildningsreformer*. Stockholm: Liber utbildningsförlaget.

*Lov om grunnskolen og den vidaregåande opplæringa av 1998 med endringar per 10. juni 1999, § 9-2.*

Lundgren, U.P. 1990. OECD-rapporten – en bakgrunn. I: Granheim, M., Lundgren, U.P. & Tiller, T. (Red.) 1990. *Utdanningskvalitet – styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole*. Oslo: Tano.

Lundgren, U.P. 2003. The Political Governing (Governance) of Education and Evaluation. I: Haug, P. & Schwandt, T.A. (Red.) 2003. *Evaluating educational reforms. Scandinavian perspectives*. Greenwich, Connecticut: Information Age Publishing.

Madaus, G.F., Scriven, M. & Stufflebeaum, D.L. 1993. Program Evaluation: A Historical Overview. I: Madaus, G.F., Scriven, M. & Stufflebeaum, D.L. (Red.). 1993. *Evaluation Models. Viewpoints on educational and human services evaluation*. Boston: Kluwer-Nijhoff Publishing.

Mintzberg, H. 1994. *The Rise and Fall of Strategic Planning: Reconceiving the Roles for Planning, Plans, Planners*. New York: Free Press.

Moe-utvalget 1997. *Nasjonalt vurderingssystem for grunnskolen*. Oslo: Kirke-, undervisnings- og forskningsdepartementet.

NOU 1995:18 *Ny lovgivning om opplæringa*. Oslo: Kirke-, undervisnings- og forskningsdepartementet.

NOU 2000:22 *Om oppgavefordelingen mellom stat, region og kommune*. Oslo: Kommunal- og regionaldepartementet.

NOU 2002:10 *Førsteklasses fra første klasse. Forslag til rammeverk for et nasjonalt kvalitetsvurderingssystem av norsk grunnsopplæring*. Oslo: Utdannings- og forskningsdepartementet.

NOU 2003:16 *I første rekke*. Oslo: Utdannings- og forskningsdepartementet.

OECD 1988. *Review of Educational Policy in Norway. Examiners Report and Questions*. Paris: OECD, Department of the Examiners.

Olsen, J.P. 1986. Foran en ny offentlig revolusjon. *Nytt Norsk Tidsskrift*, Nr. 3, 1986, 3-15.

Perrin, B. 1998. Effective Use and Misuse of Performance Measurement. *American Journal of Evaluation*, Vol. 19, No. 3, 1998, 367-379.

Perrin, B. 1999. Performance Measurement: Does the Reality Match the Rhetoric? A Rejoinder to Bernstein and Winston. *American Journal of Evaluation*, Vol. 20, No. 1, 1999, 101-111.

Phillips, D.C. 2000. *The Expanded Social Scientist's Beastiary*. Lanham: Rowman & Littlefield Publishers, Inc.

Pirsig, R.M. 1997. *Zen og kunsten å vedlikeholde en motorsykkel. En undersøkelse av verdier*. Oslo: Pax Forlag.

Rhyn, H. 2003. What's Being Done in the Name of Evaluation?: Experiences Drawn From the Recent Evaluation of Schooling Reforms in Switzerland. I: Haug, P. & Schwandt, T.A. (Red.) 2003. *Evaluating educational reforms. Scandinavian perspectives*. Greenwich, Connecticut: Information Age Publishing.

Ryen, A. 2002. *Det kvalitative forskningsintervjuet. Fra vitenskapsteori til feltarbeid*. Bergen: Fagbokforlaget.

Røvik, K.A. 1998. *Moderne organisasjoner. Trender i organisasjonstenkingen ved tusenårsskiftet*. Bergen: Fagbokforlaget.

Scriven, M. 1991a. Beyond Formative and Summative Evaluation. I: McLaughlin, M.W. & Phillips, D.C. (Red.). *Evaluation and Education: At Quarter Century*. Chicago: University of Chicago Press.

Scriven, M. 1991b. *Evaluation Thesaurus*. 4. ed. Newbury Park: Sage Publications.

Shadish, W.R., Cook, T.D. & Leviton, L.C. 1995. *Foundations of program evaluation. Theories of practice*. Newbury Park: SAGE Publications.

St.meld. nr. 30 (2003-2004) *Kultur for læring*. Oslo: Utdannings- og forskningsdepartementet.

St.meld. nr. 31 (2000-2001) *Kommune, fylke, stat – en bedre oppgavefordeling*. Oslo: Kommunal- og regionaldepartementet.

St.meld. nr. 33 (1991-1992) *Kunnskap og kyndighet. Om visse sider ved vidregående opplæring*. Oslo: Kirke-, undervisnings- og forskningsdepartementet.

St.meld. nr. 37 (1990-1991) *Om organisering og styring av utdanningssektoren*. Oslo: Kirke-, undervisnings- og forskningsdepartementet.

St.prp. nr. 1 (1997-1998) *For budsjetterminen 1998*. Oslo: Kirke-, utdannings- og forskningsdepartementet.

Telhaug, A.O. & Mediås, O.A. 2003. *Grunnskolen som nasjonsbygger. Fra statspietisme til nyliberalisme*. Oslo: Abstrakt forlag.

Thune, T. 2000. Towards a Politics of Quality. National Quality Assurance of Higher Education and Research in Britain. I: *Studies in Comparative and International Education* (vol. 6). Oslo: Universitetet i Oslo, Pedagogisk forskningsinstitutt (Rapport nr. 9, 2000).

Tiller, T. 1990. Kvalitetssikring i hverdagen – Hva med grunnleggende verdier? I: Granheim, M., Lundgren, U.P. & Tiller, T. (Red.) 1990. *Utdanningskvalitet – styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole*. Oslo: Tano.

Vanebo, J.O., Busch, T., Klausen, K.K. & Johnsen, E. 2001. Innledning – om modernisering og bokens innhold. I: Busch, T., Johnsen, E., Klausen, K.K., Vanebo, J.O. (Red.) 2001. *Modernisering av offentlig sektor. New Public Management i praksis*. Oslo: Universitetsforlaget.

Vedung, E. 1998. *Utvärdering i politik och förvaltning*. 2. utg. Lund: Studentlitteratur.

Vetlesen, A.J. 1998. *Dydsetikk i individualiseringens og globaliseringens tid. I:*

Vetlesen, A.J. (Red.) 1998. *Dydsetikk*. Oslo: Humanistisk forlag.

Weiss, C. 1998. *Evaluation*. 2. ed. Upper Saddle River, NJ: Prentice Hall

Weiler, H.N. 1988/89. Decentralisation in educational governance: an exercise in contradiction? I: Granheim, M., Kogan, M. & Lundgren, U.P. (Red.) 1990. *Evaluation as policymaking. Introducing evaluation into a national decentralised educational system*. London: Jessica Kingsley Publishers.

Winston, J.A. 1999. Performance Indicators – Promises Unmet: A Response to Perrin. *American Journal of Evaluation*, Vol. 20, No. 1, 1999, 95-99.

Winther Jørgensen, M. & Phillips, L. 1999. *Diskursanalyse som teori og metode*. Roskilde: Roskilde Universitetsforlag.

Østerud, S. 1998. Relevansen av begrepene ”validitet” og ”reliabilitet” i kvalitativ forskning. *Norsk pedagogisk tidsskrift*, nr. 3/98.

## WEB-ADRESSER

[www.bibsys.no](http://www.bibsys.no)

[www.odin.dep.no/ufd](http://www.odin.dep.no/ufd)

[www.oslo.kommune.no](http://www.oslo.kommune.no)

[www.skoleporten.no](http://www.skoleporten.no)